DOCUMENT RESUME

ED 478 170 TM 035 065

AUTHOR	Zhang, Yanwei Oliver; Yu, Feng; Nandakumar, Ratna
TITLE	The Impact of Conditional Scores on the Performance of DETECT.
PUB DATE	2003-04-00
NOTE	52p.; Paper presented at the National Council on Measurement in Education (Chicago, IL, April 22-26, 2003).
PUB TYPE	Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE	EDRS Price MF01/PC03 Plus Postage.
DESCRIPTORS	Estimation (Mathematics); *Nonparametric Statistics; *Scores; Simulation
IDENTIFIERS	Centering; *Conditionals; Covariation; *Dimensionality (Tests); Number Right Scoring

ABSTRACT

DETECT is a nonparametric, conditional covariance-based procedure to identify dimensional structure and the degree of multidimensionality of test data. The ability composite or conditional score used to estimate conditional covariance plays a significant role in the performance of DETECT. The number correct score of all items in the test (T) and the number correct score of remaining items (S), other than the two items in consideration, are two natural candidates for computing conditional covariances. However, these conditional scores produce biased estimates in finite samples. Some type of correction is required in computing the estimates of conditional covariances. This study investigated the effect of centering and/or averaging T and S as bias correction methods. This process resulted in six different estimates of conditional covariances for use in the DETECT procedure, and 72 types of test data were simulated to vary in sample size, test length, degree of multidimensionality, and distribution of items into clusters. The impact of the six estimates on the performance of DETECT were studied for three aspects: Dmax value, r ratio, and the percentage of items correctly classified into clusters. The results show that the centered conditional score S performed best. The next best index was the average of T and S with centering, followed by the average of T and S without centering. (Contains 8 tables, 11 figures, and 11 references.) (Author/SLD)

The Impact of Conditional Scores on the Performance of DETECT

Yanwei Oliver Zhang

American Institute of Certified Public Accountants, ozhang@aicpa.org

Feng Yu

Educational Testing Service

and

Ratna Nandakumar

University of Delaware

Paper presented at the annual meeting of the National Council on Measurement in Education, Chicago, April 24, 2003.

Acknowledgements: The authors wish to thank Louis Roussos and William Stout for providing the DETECT computer program.

Abstract

DETECT is a nonparametric, conditional covariance based procedure to identify dimensional structure and the degree of multidimensionality of test data. The ability composite or conditional score used to estimate conditional covariance plays a significant role in the performance of DETECT. The number correct score of all items in the test (T) and the number correct score of remaining items (S), other than the two items in consideration, are two natural candidates for computing conditional covariances. However these conditional scores produce biased estimates in finite samples. Some type of correction is required in computing the estimates of conditional covariances. This study investigated the effect of centering and/or averaging T and S as bias correction methods. This process resulted in six different estimates of conditional covariances for use in the DETECT procedure. 72 types of test data were simulated that vary in sample size, test length, degree of multidimensionality, and distribution of items into clusters. The impact of the six estimates on the performance of DETECT were studied on three aspects: Dmax value, r ratio, and the percentage of items correctly classified into clusters. The results showed that, the centered conditional score S performed the best. The next best index was the average of T and S with centering, followed by the average of T and S without centering.

DETECT (Zhang & Stout, 1999a, 1999b) is a nonparametric statistical procedure for the dimensionality assessment of binary data resulting from monotone models. It determines the simple latent dimensionality structure of a test composed of dichotomously scored items. A simple dimensionality structure of a test means that each item in the test can be classified into one and only one of the dimensionsional clusters. Given a set of items, the DETECT procedure partitions items into separate clusters so that items within the same cluster are substantively dimensional homogeneous and clusters are substantively dimensional distinct from each other. In addition to discovering the dimensional structure of a test, the DETECT procedure also reveals the seriousness of multidimensionality present in the test data. Zhang and Stout (1999b) showed that the DETECT procedure is very effective in this regard when in fact there exists a simple structure underlying a set of item response data.

The mathematical logic behind the DETECT procedure is that, for examinees with similar abilities, items measuring the same latent abilities are likely to have positive covariances. On the contrary, items measuring different latent abilities tend to have negative conditional covariances (Rosenbaum, 1988; Douglas, Kim, & Stout, 1994). Therefore, conditional covariance forms the basic building block of the DETECT procedure. The conditional score used in computing the conditional covariances of item pairs plays a critical role on the performance of DETECT procedure, and it directly affects the precision of what the procedure yields. Currently two ways are available to compute the conditional covariances of the DETECT index: (1) Kim (1994) proposed computing the covariance conditional upon the total score of remaining items other than the two items in consideration, and centered, that is, $(\widehat{\text{Cov}}_{i_1 i_2}(S) - \overline{\widehat{\text{Cov}}(S)})$; and (2) Zhang and Stout (1999b) have proposed taking the average of two conditional covariances: one based on the total score of the test $\widehat{\text{Cov}}_{i_1 i_2}(T)$ and the other based on the total score of the remaining items on the test, $\widehat{\text{Cov}}_{i_1 i_2}(S)$. There are clearly other ways of choosing the conditional score in computing conditional covariances. The choice of the conditional score could greatly impact the performance of the DETECT

index. To date, no extensive study has been done to investigate different ways of computing conditional covariances and how they benefit the DETECT procedure the most.

The purpose of the present study is to investigate the effect of different conditional scores on the performance of the DETECT procedure through simulated multidimensional data. Three conditional scores will be considered along with the option of centering (centered and uncentered), resulting in six different ways of computing the DETECT index: (1) condition on total number right score uncentered, (2) condition on the number right on remaining items of the test uncentered, (3) Average 1 and 2, (4) condition on total number right score centered, (5) condition on the number right on remaining items of the test centered, (6) average of 4 and 5.

The Theoretical DETECT Index

DETECT is an extremely powerful technique that is based on a strong theory of conditional covariances and utilizes a genetic algorithm to arrive at the partition of items that quantifies the maximum degree of multidimensionality present in the given test data. The theoretical computation of DETECT index is described briefly as follow (for details, see Zhang & Stout, 1999b).

Let N denote the number of dichotomous items of a test. Let $\mathcal{P} = A_1, A_2, \cdots, A_k$ denote a partition of the N test items into k clusters. The theoretical DETECT index, which gives the degree of multidimensionality of the partition $\mathcal{P}$ is defined as,

$$D(\mathcal{P}, \Theta_T) = \frac{2}{N(N-1)} \sum_{1 \le i \le j \le N} \delta_{ij}(\mathcal{P}) E[\mathrm{Cov}(X_i, X_j | \Theta_T = \theta)], \qquad (1)$$

where, Θ_T is the test composite, X_i, X_j are scores on items i and j, and

$$\delta_{i_1 i_2} = \begin{cases} 1 & \text{if items } i \text{ and } j \text{ are in the same cluster of } \mathcal{P} \\ -1 & \text{otherwise} \end{cases} \qquad (2)$$

$D(\mathcal{P}, \Theta_T)$ measures the amount of multidimensionality present for the given partition $\mathcal{P}$. Obviously, there are numerous ways to partition items of a test into clusters, and each partition produces a value of $D(\mathcal{P}, \Theta_T)$. The partition $\mathcal{P}^*$ which produces the maximum value among all possible $D(\mathcal{P}, \Theta_T)$'s, is treated as the optimal simple dimensionality structure of the test, and $D_{max}(\mathcal{P})$ associated with $\mathcal{P}^*$, is treated as the maximum amount of multidimensionality present in the test data. For example, for a purely unidimensional test, the optimal dimensionality structure of the test is that all the items will be partitioned into one single cluster, and the value of $D_{max}(\mathcal{P})$ of the test will be close to 0. It has been shown by Zhang and Stout (1999b) that when there is a true simple structure underlying test data, $D(\mathcal{P}, \Theta_T)$ will be maximized only for the correct partition.

In order to determine if the partition $\mathcal{P}$ that maximized the DETECT index $D(\mathcal{P}, \Theta_T)$ is indeed the correct structure of the test, the following ratio can be useful:

$$r = \frac{D_{max}(\mathcal{P}, \Theta_T)}{D^*(\mathcal{P}, \Theta_T)} \tag{3}$$

where

$$D^*(\mathcal{P}, \Theta_T) = \frac{2}{N(N-1)} \sum_{1 \leq i \leq j \leq N} |E[\mathrm{Cov}(X_i, X_j | \Theta_T = \theta)]|, \tag{4}$$

When there is a simple structure underlying test data, the ratio r is close to 1. The extent to which r differs from 1 is indicative of the degree to which the structure of the test deviates from the simple structure.

Estmation of the DETECT Index

To estimate $E[\mathrm{Cov}(X_i, X_j | \Theta_T = \theta)]$, a statistic, in place of the latent variable (Θ_T), is needed as the conditional score. As discussed by Zhang and Stout (1996, 1999), there are

two natural estimators of $E[\text{Cov}(X_i, X_j|\Theta_T = \theta)]$:

$$\widehat{\text{Cov}}_{ij}(T) = \sum_{m=0}^{N} \frac{J_m}{J} \widehat{\text{Cov}}(X_i, X_j|T = m), \tag{5}$$

where the conditional score $T = \sum_{l=1}^{N} X_l$ is the total score of all test items; J is the total number of examinees; and J_m is the number of examinees in subgroup m with the total score $T = m$. $\widehat{\text{Cov}}(X_i, X_j|T = m)$ is the sample conditional covariance for examinees in the subgroup m. The other is the estimator based on the total score of remaining items given by,

$$\widehat{\text{Cov}}_{ij}(S) = \sum_{m=0}^{N-2} \frac{J_m}{J} \widehat{\text{Cov}}(X_i, X_j|S = m), \tag{6}$$

where the score $S = \sum_{l=1, l\neq i,j}^{N} X_l$ is the total score of the remaining items, other than items i and j. J_m is the number of examinees in subgroup m with the conditional score $S = m$, and $\widehat{\text{Cov}}(X_i, X_j|S = m)$ is the sample conditional covariance for examinees in the subgroup m.

When a test is unidimensional, $\widehat{\text{Cov}}_{ij}(T)$ tends to be negative because items X_i and X_j are part of T. Therefore, $\widehat{\text{Cov}}_{ij}(T)$ as an estimator of $E[\text{Cov}(X_i, X_j|\Theta_T = \theta)]$ results in a negative bias (Junker, 1993; Zhang and Stout, 1999a). $\widehat{\text{Cov}}_{ij}(S)$, on the other hand, tends to be positive and results in a positive bias (Rosenbaum, 1984; Holland and Rosenbaum, 1986; Zhang and Stout, 1999a).

In the original DETECT index Kim (1994) proposed $\widehat{\text{Cov}}_{ij}(S)$ as an estimator of $E[\text{Cov}(X_i, X_j|\Theta_T = \theta)]$, and further used a correction for the positive bias resulting in the following index:

$$D_k(\mathcal{P}) = \frac{2}{N(N-1)} \sum_{1\leq i\leq j\leq N} \delta_{ij}(\mathcal{P})[\widehat{\text{Cov}}_{ij}(S) - \overline{\text{Cov}(S)}] \tag{7}$$

where δ_{ij} is defined in Equation (2) and $\overline{\text{Cov}(S)}$ is the average of $\widehat{\text{Cov}}_{ij}(S)$ over all $N(N-1)/2$ item pairs. The average $\overline{\text{Cov}(S)}$ is subtracted from each $\widehat{\text{Cov}}_{ij}(S)$ to correct for the positive bias in the unidimensional case.

Since $\widehat{\mathrm{Cov}}_{ij}(T)$ tends to have a negative bias and $\widehat{\mathrm{Cov}}_{ij}(S)$ tends to have a positive bias as estimators of $E[\mathrm{Cov}(X_i, X_j|\Theta_T = \theta)]$ in the unidimensional case, Zhang and Stout (1999b) proposed an average of these two estimates resulting in the following index for DETECT:

$$D_{ZS}(\mathcal{P}) = \frac{2}{N(N-1)} \sum_{1 \le i \le j \le N} \delta_{ij}(\mathcal{P}) \widehat{\mathrm{Cov}}^*_{ij} \tag{8}$$

where

$$\widehat{\mathrm{Cov}}^*_{ij} = \frac{1}{2}[\widehat{\mathrm{Cov}}_{ij}(S) + \widehat{\mathrm{Cov}}_{ij}(T)]. \tag{9}$$

Zhang and Stout (1999b)'s rationale for suggesting Equation (9) is purely theoretical in nature. Zhang and Stout (1999b) recommend $D_{ZS}(\mathcal{P})$ based on their results of a small scale simulation study.

Clearly there are other possibilities for estimating $E[\mathrm{Cov}(X_i, X_j|\Theta_T = \theta)]$. For example, the total score T with (or without) the correction for the negative bias, or further correction for the average covariance, $\widehat{\mathrm{Cov}}^*_{ij}$ could improve the performance of DETECT. Six different estimates of $E[\mathrm{Cov}(X_i, X_j|\Theta_T = \theta)]$ are considered in this study for comparison purposes. The three main estimates are: 1) $\widehat{\mathrm{Cov}}_{ij}(S)$, 2) $\widehat{\mathrm{Cov}}_{ij}(T)$, 3) $\widehat{\mathrm{Cov}}^*_{ij}$. As a correction for the bias all three estimates are also centered resulting in six indices as follows.

$$D_1(\mathcal{P}) = \frac{2}{N(N-1)} \sum_{1 \le i \le j \le N} \delta_{ij}(\mathcal{P})[\widehat{\mathrm{Cov}}_{ij}(T)] \tag{10}$$

$$D_2(\mathcal{P}) = \frac{2}{N(N-1)} \sum_{1 \le i \le j \le N} \delta_{ij}(\mathcal{P})[\widehat{\mathrm{Cov}}_{ij}(S)] \tag{11}$$

$$D_3(\mathcal{P}) = \frac{2}{N(N-1)} \sum_{1 \le i \le j \le N} \delta_{ij}(\mathcal{P})\widehat{\mathrm{Cov}}^*_{ij} \tag{12}$$

$$D_4(\mathcal{P}) = \frac{2}{N(N-1)} \sum_{1 \le i \le j \le N} \delta_{ij}(\mathcal{P})[\widehat{\mathrm{Cov}}_{ij}(T) - \overline{\widehat{\mathrm{Cov}}(T)}] \tag{13}$$

$$D_5(\mathcal{P}) = \frac{2}{N(N-1)} \sum_{1 \le i \le j \le N} \delta_{ij}(\mathcal{P})[\widehat{\mathrm{Cov}}_{ij}(S) - \overline{\widehat{\mathrm{Cov}}(S)}] \tag{14}$$

$$D_6(\mathcal{P}) \;=\; \frac{2}{N(N-1)} \sum_{1\le i\le j\le N} \delta_{ij}(\mathcal{P})[\widehat{\mathrm{Cov}}^{*}_{ij} - \overline{\widehat{\mathrm{Cov}}^{*}}] \tag{15}$$

where δ_{ij}, $\mathrm{Cov}(S)$, $\mathrm{Cov}(T)$, $\widehat{\mathrm{Cov}}^{*}_{ij}$, and $\overline{\widehat{\mathrm{Cov}}(S)}$ are as defined before. $\overline{\widehat{\mathrm{Cov}}(T)}$ is the average of $\widehat{\mathrm{Cov}}_{ij}(T)$ over all $N(N-1)/2$ item pairs; and $\overline{\widehat{\mathrm{Cov}}^{*}}$ is the average of $\widehat{\mathrm{Cov}}^{*}_{ij}$ over all $N(N-1)/2$ item pairs. The index D_5 is same as D_k, and the index D_3 is the same as D_{ZS}. All the six indices are studied and compared in simulated settings of one and two-dimensional tests for their ability to correctly classify items into different clusters and to assess the degree of multidimensionality present.

In order to obtain the optimal simple dimensional structure of a test, a Genetic Algorithm (GA) is built in the DETECT procedure to correctly classify items into different dimensional clusters (Zhang & Stout, 1996b). GA is an optimization tool to iteratively change items' cluster memberships until $D(\mathcal{P}^{*})$ is reached.

Simulation of Two-dimensional Composite Test Data

A two-dimensional composite test measures two latent abilities. Each item in the test is driven by two latent abilities (θ_1, θ_2) simultaneously. Each item measures a weighted linear combination of θ_1 and θ_2, which is called a composite. The direction of an item composite is determined by its angle value in the (θ_1, θ_2) plane, which is another type of item representation used in NOHARM (a non-linear factor analysis program in McDonald's non-linear factor analysis procedure). The angle of an item composite is defined as the ratio of its discrimination parameter on θ_2 (a_2) to its discrimination parameter on θ_1 (a_1). Figure 1 illustrates the basic idea of the angle value type item representation. Line l_i represents a composite of item i. When α is less than 45^o, the composite relies more heavily on θ_1 than on θ_2; when α is larger than 45^o, the composite relies more heavily on θ_2 than on θ_1. In other words, as α increases, latent ability θ_2 contributes more to the composite. A test

is said to be essentially unidimensional if all item composites in a test lie almost in the same direction (within a narrow fan). Items within distinct narrow fans (clusters) measure different composite abilities.

Test data with a two-cluster structure (two distinct composite abilities) were simulated in the current study. Figures 2 shows an example of a test with two item clusters. Items within each cluster are believed to measure the same latent ability composite, and the clusters represent distinct latent ability composites. The smaller the angles between the clusters in a test, the less the degree of multidimensionality the test appears. Higher the percentage of items within a cluster, less the degree of multidimensionality.

The simulated angle values of the clusters used in the present study are illustrated in Table 1. As can be seen in the fist two columns of Table 1, there are six combination of angles where the angular difference between item clusters ranges from 0° to 90°, denoting the increase in the degree of multidimensionality. For each angle combination, three types of distribution of test items are considered: all test items are in Cluster 1; two-thirds of items in Cluster1 and one-third in Cluster2; and half of items Cluster1 and half in Cluster2. For example, for 30-item test: all 30 items are in Cluster1 and 0 in Cluster2; 20 items in Cluster1 and 10 in Cluster2; 15 items in Cluster1 and 15 in Cluster2. Unidimensionality results when all test items are included in one cluster; and also when the angle-difference between the clusters is 0°. The shaded rows of Table 1 indicate unidimensional tests.

In order to make the simulated data as realistic as possible, estimated item parameters from the National Assessment of Educational Progress test data (The 1992 NAEP Technical Report) were used in this study. Table 2 gives a summary of descriptive statistics of item parameters. For a two-dimensional composite item i with angle value α_i, its item parameters were defined in the following manner (Kim, 1994):

A set of unidimensional item parameter estimates a_i, b_i, and c_i was randomly selected from the 1992 NAEP item pool (collection of item parameter estimates from the report). Using these estimated unidimensional parameters, two-dimensional parameters were computed. The two-dimensional discrimination parameters were defined as

$$a_{1i} = a_i \cos(\alpha_i) \qquad a_{2i} = a_i \sin(\alpha_i).$$

Difficulty and guessing parameters for a dichotomous item were defined as

$$b_{1i} = b_{2i} = b_i, \quad c_i = c_i.$$

The amount of the multidimensionality in a two-dimensional composite test was determined by the angle between the clusters, the correlation between abilities (which was fixed to 0.3 in the current study), and the distribution of test items in the two clusters.

Each examinee's abilities θ_1 and θ_2 are randomly generated from a bivariate normal distribution with the correlation coefficient between the abilities fixed at 0.3. Two test lengths: 30 and 60 were considered in the present study. In all there are 36 different combination of tests (2 test lengths x 6 angle-combinations x 3 item distributions to clustrs) denoting different degrees of multidimensionality in items. Each of these 36 combinations is crossed with two sample sizes (500 and 1000) producing 72 types of test data. In each case dichotomous item responses were generated according to the three-parameter logistic model with two compensatory abilities (Reckase & McKinley, 1983) given by

$$P_i(\theta_{1j}, \theta_{2j}) = c_i + \frac{1 - c_i}{1 + exp[-1.7[a_{1i}(\theta_{1j} - b_{1i}) + a_{2i}(\theta_{2j} - b_{2i})]]}, \tag{16}$$

where $P_i(\theta_{1j}, \theta_{2j})$ is the probability of correct response to the dichotomous item i by an examinee j with ability $(\theta_{1j}, \theta_{2j})$. a_{1i} is the discrimination parameter of the dichotomous item i on θ_1 and a_{2i} is the discrimination parameter of the item i on θ_2. Similarly, b_{1i} is the difficulty parameter of item i on θ_1; and b_{2i} is the difficulty parameter of item i on θ_2. c_i is the guessing parameter of the item i.

For each simulated examinee, response probability for each item was computed using the above equation. If the computed probability was greater than the uniform random variable generated from the interval (0,1), then the item was considered answered correctly and a score of 1 was assigned. Otherwise a score of 0 was assigned.

For each test data (72 in all), the DETECT procedure was implemented six times and the following information was recorded for each of the six indices, D_1 through D_6:

- The percentage of items correctly classified into intended clusters.
- D_{max} – the maximum value of $D(\mathcal{P})$ associated with the best simple structure solution available for given data.
- $r = \frac{D_{max}(\mathcal{P})}{D^*(\mathcal{P})}$ – the degree of divergence from simple structure. The degree to which r is further from 1 indicates the degree of divergence, of the dimensional structure of given data, from the simple structure solution.

This procedure was replicated 100 times and results were averaged and tabulated.

Results

Tables 3 and 4 show results for the percentage of items correctly classified into intended clusters. Table 3 shows results for 30 items and Table 4 for 60 items. Cell values are the mean percentages, over 100 replications, of items correctly classified into clusters along with the standard deviations. In each table there are six panels, one panel for each angular difference. Each panel contains percentages of items correctly classified in the three different item distributions for both sample sizes (across columns) for all six indices (across rows). These results are plotted in Figures 3 and 4. Figures 3 shows graphs for 30 items and Figure

4 for 60 items. Each of Figures 3 and 4 have four graphs (two clusters crossed with two sample sizes).[1] Some general expected trends can be observed from these figures. It can be seen from Figures 3 and 4 (3a through 3d; 4a through 4d) that, as the angular difference between the clusters decrease from 90° to 10°, the percentage of items correctly classified into clusters decreases from near 100% to about 50%. As the number of examinees increases, the percentage correct increases, especially for the angular differences smaller than 50°. This is true for both test lengths. There is a slight increase in the percentage correct when items are equally distributed into clusters (15/15 and 30/30 items) than unequal distribution (20/10 and 40/20).

Regarding the comparison of six methods with respect to the percentage of items correctly classified, the overall performance of D_1 and D_2 is not consistent. For example, D_2 is high only when the angular difference between the clusters is large such as 90° and 70°, coupled with a higher percentage of items within a cluster. In other cases D_2 performed badly. Whereas the performance of D_3 through D_6 show a consistent pattern across conditions. Average percentages of items correctly classified collapsed over sample sizes and item distributions are plotted in Figures 9a and 9b. Performance of D_1 and D_2 is not further considered as they are unreliable. Careful examination of the performance of the rest of the four indices D_3 through D_6 (for all item distributions, different angular differences, and sample sizes) reveals that, overall, D_4(conditioning on N items with centering) is slightly better than D_6 average with centering), which is slightly better than D_5 (conditioning on $N-2$ items with centering).

Results for D_{max} are listed in Tables 5 and 6 for 30 and 60 items respectively. These Tables are organized similar to Tables 3 and 4. In that there are six panels, one for each angular difference. The cell values of the table give mean D_{max} values along with their

[1]The percent correct for unidimensional cases are not graphed as there is only one cluster.

standard deviations over 100 replications. The D_{max} values are plotted in Figures 5 and 6 for 30 and 60 items respectively. Each figure contains six graphs (3 item distributions by 2 sample sizes).

Unidimensionality results when all test items are in one cluster (graphs 5a, 5b, 6a, and 6b), irrespective of the angular differences between clusters. In all unidimensional graphs, as expected, the plots are straight lines. That is, for unidimensional tests, the direction of the best measurement has no effect on D_{max} value, as it should be. It can also be seen that D_{max} values decreases as the test size and sample size increase. Since the test size and sample size influence conditional covariances, they in turn affect the values of D_{max}. Average D_{max} values collapsed over sample sizes and item distributions are plotted in Figures 10a and 10b. From these figures it is evident that, these averages are all small (less than .20), and D_3 and D_4 are lowest, followed by D_5 and D_6, although these differences are small.

The rest of the graphs in Figures 5 and 6 are for two-dimensional data. In two-dimensional test data, the emphasis is on power. That is, one expects higher D_{max} values associated with test data reflecting higher degree of multidimensionality; and for these indices to be sensitive to the degree of multidimensionality present in test data. It can be seen from these figures that as the degree of multidimensionality decreases from near simple structure sceneriio to a single unidimensional cluster, the D_{max} values range from high .90 to below .20. D_{max} values are higher when items are equally distributed into clusters (30/30 or 15/15) than for unequal distributions (40/20 or 20/10). They are also higher for 30 item test data than 60 item test data. Average D_{max} values collapsed over sample sizes and item distributions are plotted in Figures 10c and 10d. Comparison of D_2 through D_6 shows that the general pattern appears to be favoring D_5, which has highest D_{max} values, followed by D_6 and D_3, and then D_4, although the differences are small.

Results of r values are listed in Tables 7 and 8 for 30 and 60 items respectively.

These tables are organized similar to Tables 3 to 6. The cell values give mean r values and their standard deviations over 100 replications. These results are plotted in Figures 7and 8 for 30 and 60 items respectively. Each of these figures have four panels corresponding to multidimensional test data. One expects to find r values near 1 for simple structure test data (with angular difference of 90° between the clusters and equal distribution of items into clusters). It can be seen that r values for simple structure test data are the highest (about .90) and gradually decrease as the angular difference decreases from 90° to 10°. Moreover, the r values increase as the sample size increases; higher for equally distributed clusters (15/15 and 30/30). r values are higher for 30 item test data than 60 item test data. Average r values collapsed over sample sizes and item distributions are plotted in Figures 11a and 11b. Comparing D_3 through D_6 it can be seen that all indices seem to be functioning equally well for 90° and 70°. At 50° or below D_5 seem to be performing slightly better than others, D_3 and D_6 are about the same followed by D_4.

Summary and Discussion

The conditional score used in estimating the conditional covariance could play a significant role on the performance of DETECT as a procedure to identify the simple dimensional structure underlying test data and to quantify the degree of multidimensionality present in data. This study is a pilot to investigate the performance of DETECT procedure with respect to different conditional scores. There are two natural ways for estimating the conditional covariance: conditioned on the total test score (T), conditioned on the score of remaining items (S). As discussed earlier, both of these estimated covariances are biased estimates, in opposite directions, of the true conditional covariance. Kim (1994) suggested the use of **centered** S, where centering served as a correction for the positive bias in S. Zhang and Stout (1999b) suggested taking the average of the conditional covariances, S and T so that the bias gets canceled out. Naturally, there are other ways of obtaining estimated

conditional covariances. In this study two types of bias correction were considered: averaging the conditional covariances T and S, and centering the indices (centered versus uncentered), resulting in six different estimates of conditional covariances.

The Performance of the DETECT procedure was compared using all six indices D_1 through D_6 with respect to: (a) the percentage of items correctly classified, (b) D_{max} values, and (c) r values, for varied test length, item distributions, and sample sizes. Since D_1 and D_2 showed unreliable and inconsistent results with respect to the percentage of items correctly classified into clusters, these were eliminated for further consideration. Comparison of D_3, D_4, D_5, and D_6 showed that in ideal situations, such as clear simple structure solution with an angular difference of 90° to 70°, all four indices performed similarly. However, in other cases, D_4 was best with respect to the percentage of items correctly classified into clusters, followed by D_6 and by D_5. Whereas D_5 was better than other indices with respect to D_{max} and r values. Based on this study it can be concluded that, taken as a whole, D_5 (conditioned on S and centering) performed best. The next best index was D_6 (average of T and S with centering), followed by D_3 (average of T and S without centering), although the difference between D_6 and D_3 is trivially small.

This study is limited. Although we have investigated six different correction procedures for the bias in estimating the conditional covariance, there are more ways to correct for this bias. For example, based on theory, conditional covariance based on T is negatively biased. Hence it makes sense to center it by adding the mean conditional covariance instead of subtracting, as done in this study. Whereas for the average conditional covariance, $\widehat{\text{Cov}}^*_{ij}$, the theoretical argument suggests the bias can cancel out by averaging. However, it is not known, if the magnitude of bias is same in both directions. Perhaps one could shift the estimated covariance positively and negatively to see what works best. Although D_5 is the recommended procedure based on this limited study, a more detailed follow-up study is

needed to confirm this conclusion.

Educational tests are increasingly becoming intentionally multidimensional. Modeling such data by the test composite, a composite of latent traits, is a convenient and useful approach for determining the dimensional structure of the test. In this regard, DETECT is an extremely effective procedure to detect the multidimensional structure of data, and to classify items into dimensional clusters. Given the importance of the DETECT procedure for the assessment of educational test data, more studies on the usefulness of the procedure is important.

References

Kim, H. R. (1994). *New techniques for the dimensionality assessment of standardized test data.* Unpublished doctoral dissertation, University of Illinois at Urbana-Champaign.

Holland, P. W., & Rosenbaum, P. R. (1986). Conditional association and unidimensionality in monotone latent variable models. *The Annals of Statistics, 14*, 1523-1543.

Junker, B. (1993). Conditional association, essential independence and monotone unidimensional item response models. *The Annals of Statistics, 21*, 1359-1378.

McDonald, P. R. (1962). A general approach to nonlinear factor analysis. *Psychometrika, 4*, 397-415.

McDonald, P. R., & Ahlawat, K. S. (1974). Different factors in binary data. *British Journal of Mathematical and Statistical Psychology, 27*, 82-99.

Reckase, M. D., & Mckinley, R. L. (1983). *The definition of difficulty and discrimination for multidimensional item response theory models.* Paper presented at the Annual Meeting of American Educational Research Association, Montreal.

Rosenbaum, P. R. (1984). Testing the conditional independence and monotonicity assumptions of item response theory. *Psychometrika, 49*, 425-435.

Roussos, L. A., Stout, W. F., & Marden, J. I. (1993). *Dimensional and structural analysis of standardized tests using DIMTEST with hierarchical cluster analysis.* Paper presented at the annual NCME meeting, Atlanta, GA.

Stout, W., Habing, B., Douglas, J., Kim, H. R., Roussos, L., & Zhang, J. (1996). Conditional covariance-based nonparametric multidimensionality assessment. *Applied Psychologi-*

cal Measurement, 20, 331-354.

Zhang, J., & Stout, W. F.(1999a). Conditional Covariance structure of generalized compensatory multidimensional items. *Psychometrika, 64*, 129-152.

Zhang, J., & Stout, W. F.(1999b). The theoretical DETECT index of dimensionality and its application to approximate simple structure. *Psychometrika, 64*, 213-249.

Table 1: Angles and Items in Each Cluster (α is the angle between items in Cluster1 and θ_1-direction, β is the angle between items in Cluster2 and θ_1-direction)

Angles	Angle between Clusters	30 items		60 items	
		Cluster1	*Cluster 2*	*Cluster 1*	*Cluster 2*
$\alpha=90^0$, $\beta=0^0$	90	30	0	60	0
		20	10	40	20
		15	15	30	30
$\alpha=80^0$, $\beta=10^0$	70	30	0	60	0
		20	10	40	20
		15	15	30	30
$\alpha=70^0$, $\beta=20^0$	50	30	0	60	0
		20	10	40	20
		15	15	30	30
$\alpha=60^0$, $\beta=30^0$	30	30	0	60	0
		20	10	40	20
		15	15	30	30
$\alpha=50^0$, $\beta=40^0$	10	30	0	60	0
		20	10	40	20
		15	15	30	30
$\alpha=45^0$, $\beta=45^0$	0	30	0	60	0
		20	10	40	20
		15	15	30	30

* Shaded cells represent unidimensional tests.

Table 2: Descriptive Statistics of Item Parameters

	a	b	c
Mean	1.007	0.282	0.202
Std.	0.503	1.345	0.100
Max.	2.615	5.093	0.465
Min.	0.281	-5.565	0.000

Table 3: Percentage of Items Correctly Classified into Clusters for 30 Items (Standard Deviations Are Presented in Smaller Font)

Type of Conditional Score	Number of Items in Each Cluster					
	30-0		20-10		15-15	
	N=500	*N=1000*	*N=500*	*N=1000*	*N=500*	*N=1000*
Angle between Clusters = 90 (α=90^0, β=0^0)						
All N items W/O Centering	100	100	74.57	75.60	84.20	86.23
	0	0	4.51	3.87	4.61	3.63
N-2 items W/O Centering	100	100	98.47	99.10	98.23	99.43
	0	0	2.39	1.63	2.34	1.34
Average W/O Centering	100	100	91.40	92.87	94.57	96.57
	0	0	4.67	3.11	3.93	2.90
All N items W Centering	100	100	94.30	96.47	96.60	98.17
	0	0	4.19	2.80	3.32	2.39
N-2 items W Centering	100	100	92.40	93.93	94.10	95.50
	0	0	3.70	2.86	3.38	2.70
Average W Centering	100	100	93.40	95.27	96.00	97.53
	0	0	4.32	3.08	3.61	2.58
Angle between Clusters = 70 (α=80^0, β=10^0)						
All N items W/O Centering	100	100	70.13	72.53	75.63	78.20
	0	0	4.97	3.76	4.44	3.46
N-2 items W/O Centering	100	100	86.60	83.57	85.13	86.83
	0	0	14.30	15.79	19.46	20.24
Average W/O Centering	100	100	82.67	86.00	87.10	89.63
	0	0	6.60	3.82	5.72	4.44
All N items W Centering	100	100	85.23	89.03	88.43	91.37
	0	0	6.95	4.62	6.15	5.08
N-2 items W Centering	100	100	83.40	87.87	87.03	89.47
	0	0	6.62	3.43	4.44	3.69
Average W Centering	100	100	84.70	88.13	88.50	91.33
	0	0	5.99	4.00	5.39	4.74
Angle between Clusters = 50 (α=70^0, β=20^0)						
All N items W/O Centering	100	100	59.63	65.77	65.67	70.30
	0	0	6.93	5.68	6.10	4.24
N-2 items W/O Centering	100	100	65.50	65.93	51.37	50.77
	0	0	2.61	1.54	2.33	1.41
Average W/O Centering	100	100	66.63	77.00	74.03	79.37
	0	0	11.45	9.27	9.75	6.75
All N items W Centering	100	100	69.90	79.10	74.23	79.57
	0	0	11.26	7.37	9.16	7.53
N-2 items W Centering	100	100	61.53	72.27	72.43	75.77
	0	0	13.83	13.56	11.58	12.69
Average W Centering	100	100	68.33	78.23	75.53	81.10
	0	0	11.42	8.91	8.99	7.79

(To be continued)

Table 3 (Continued): Percentage of Items Correctly Classified into Clusters for 30 Items (Standard Deviations Are Presented in Smaller Font)

Type of Conditional Score	Number of Items in Each Cluster					
	30-0		20-10		15-15	
	N=500	*N=1000*	*N=500*	*N=1000*	*N=500*	*N=1000*
Angle between Clusters = 30 ($\alpha=90^0$, $\beta=0^0$)						
All N items W/O Centering	100 0	100 0	46.30 6.46	50.47 7.89	48.43 6.69	51.90 8.21
N-2 items W/O Centering	100 0	100 0	65.27 2.93	65.73 1.78	51.63 2.86	51.00 1.80
Average W/O Centering	100 0	100 0	48.20 7.79	52.77 8.94	50.53 5.78	51.67 7.15
All N items W Centering	100 0	100 0	50.97 7.75	57.23 9.13	51.57 6.52	54.50 8.75
N-2 items W Centering	100 0	100 0	49.53 7.42	51.57 7.03	50.60 5.87	49.93 5.19
Average W Centering	100 0	100 0	48.57 6.95	53.13 8.71	50.63 6.15	51.90 6.80
Angle between Clusters = 10 ($\alpha=50^0$, $\beta=40^0$)						
All N items W/O Centering	100 0	100 0	42.73 4.86	42.90 5.10	43.10 5.32	43.27 4.44
N-2 items W/O Centering	100 0	100 0	64.80 2.97	65.40 2.21	51.40 1.91	50.93 1.58
Average W/O Centering	100 0	100 0	47.47 6.85	49.10 5.56	47.57 5.74	47.80 4.42
All N items W Centering	100 0	100 0	49.57 6.28	50.97 6.95	47.20 5.29	48.40 4.70
N-2 items W Centering	100 0	100 0	49.30 6.25	49.83 5.79	48.83 4.60	48.47 3.80
Average W Centering	100 0	100 0	48.57 6.62	49.97 5.71	48.77 5.76	48.77 3.87
Angle between Clusters = 0 ($\alpha=45^0$, $\beta=45^0$)						
All N items W/O Centering	100 0	100 0	100 0	100 0	100 0	100 0
N-2 items W/O Centering	100 0	100 0	100 0	100 0	100 0	100 0
Average W/O Centering	100 0	100 0	100 0	100 0	100 0	100 0
All N items W Centering	100 0	100 0	100 0	100 0	100 0	100 0
N-2 items W Centering	100 0	100 0	100 0	100 0	100 0	100 0
Average W Centering	100 0	100 0	100 0	100 0	100 0	100 0

Table 4: Percentage of Items Correctly Classified into Clusters for 60 Items (Standard Deviations Are Presented in Smaller Font)

Type of Conditional Score	Number of Items in Each Cluster					
	60-0		40-20		30-30	
	N=500	*N=1000*	*N=500*	*N=1000*	*N=500*	*N=1000*
Angle between Clusters = 90 ($\alpha=90^0$, $\beta=0^0$)						
All N items W/O Centering	100	100	75.00	78.22	90.42	92.43
	0	0	4.14	3.25	3.15	2.62
N-2 items W/O Centering	100	100	97.93	99.13	98.55	99.55
	0	0	1.88	1.10	1.33	0.82
Average W/O Centering	100	100	91.12	93.40	96.70	98.25
	0	0	4.63	2.73	2.18	1.56
All N items W Centering	100	100	90.58	93.50	97.72	98.87
	0	0	5.41	3.08	1.92	1.36
N-2 items W Centering	100	100	91.90	94.13	95.85	97.37
	0	0	3.00	1.92	2.27	1.66
Average W Centering	100	100	92.45	94.57	97.00	98.43
	0	0	3.98	2.51	2.22	1.46
Angle between Clusters = 70 ($\alpha=80^0$, $\beta=10^0$)						
All N items W/O Centering	100	100	68.57	73.12	80.08	83.60
	0	0	3.82	3.34	3.59	2.85
N-2 items W/O Centering	100	100	93.63	97.68	95.48	97.88
	0	0	5.81	1.83	2.72	1.62
Average W/O Centering	100	100	82.27	87.20	89.80	93.10
	0	0	5.23	3.59	4.27	3.29
All N items W Centering	100	100	81.08	86.97	91.42	95.00
	0	0	7.03	3.98	4.58	2.92
N-2 items W Centering	100	100	84.67	88.30	89.05	91.77
	0	0	4.58	2.64	3.34	2.83
Average W Centering	100	100	83.82	88.48	90.70	93.92
	0	0	5.05	3.72	4.18	3.09
Angle between Clusters = 50 ($\alpha=70^0$, $\beta=20^0$)						
All N items W/O Centering	100	100	56.68	64.78	65.88	72.90
	0	0	7.03	4.50	5.63	4.25
N-2 items W/O Centering	100	100	69.73	66.60	63.10	56.77
	0	0	7.99	3.35	17.88	15.60
Average W/O Centering	100	100	67.15	76.55	74.00	83.53
	0	0	8.58	5.53	6.99	4.40
All N items W Centering	100	100	63.95	76.12	74.28	84.20
	0	0	9.66	6.90	8.29	5.06
N-2 items W Centering	100	100	69.65	78.33	76.15	82.37
	0	0	8.68	6.88	7.06	6.17
Average W Centering	100	100	67.73	78.43	74.37	84.03
	0	0	8.66	5.17	7.17	4.53

(To be continued)

Table 4 (Continued): Percentage of Items Correctly Classified into Clusters for 60 Items (Standard Deviations Are Presented in Smaller Font)

Type of Conditional Score	Number of Items in Each Cluster					
	60-0		40-20		30-30	
	N=500	*N=1000*	*N=500*	*N=1000*	*N=500*	*N=1000*
Angle between Clusters = 30 ($\alpha=90^0$, $\beta=0^0$)						
All N items W/O Centering	100	100	43.72	48.27	45.90	51.48
	0	0	4.93	5.74	6.13	7.62
N-2 items W/O Centering	100	100	65.13	66.20	51.85	51.00
	0	0	3.05	1.54	2.28	1.34
Average W/O Centering	100	100	49.13	55.20	48.30	54.02
	0	0	7.70	7.31	5.99	8.62
All N items W Centering	100	100	48.38	54.68	48.77	54.47
	0	0	7.52	7.44	6.81	7.92
N-2 items W Centering	100	100	53.20	58.48	49.43	50.95
	0	0	7.50	5.27	6.28	6.93
Average W Centering	100	100	49.08	56.03	48.60	53.95
	0	0	7.48	7.07	6.79	9.49
Angle between Clusters = 10 ($\alpha=50^0$, $\beta=40^0$)						
All N items W/O Centering	100	100	41.22	41.72	41.25	41.25
	0	0	4.19	4.09	3.60	3.44
N-2 items W/O Centering	100	100	64.95	65.97	51.62	50.90
	0	0	3.02	1.73	2.37	1.17
Average W/O Centering	100	100	45.45	50.43	44.17	45.15
	0	0	5.64	6.08	4.36	4.24
All N items W Centering	100	100	45.08	50.62	44.90	45.77
	0	0	5.95	6.23	3.41	3.76
N-2 items W Centering	100	100	50.17	54.00	45.80	47.30
	0	0	6.61	4.67	4.81	3.73
Average W Centering	100	100	45.98	51.70	43.97	45.47
	0	0	6.17	5.61	4.07	4.07
Angle between Clusters = 0 ($\alpha=45^0$, $\beta=45^0$)						
All N items W/O Centering	100	100	100	100	100	100
	0	0	0	0	0	0
N-2 items W/O Centering	100	100	100	100	100	100
	0	0	0	0	0	0
Average W/O Centering	100	100	100	100	100	100
	0	0	0	0	0	0
All N items W Centering	100	100	100	100	100	100
	0	0	0	0	0	0
N-2 items W Centering	100	100	100	100	100	100
	0	0	0	0	0	0
Average W Centering	100	100	100	100	100	100
	0	0	0	0	0	0

Table 5: *D-max* Value for 30 Items (Standard Deviations Are Presented in Smaller Font)

Type of Conditional Score	Number of Items in Each Cluster					
	30-0		20-10		15-15	
	N=500	*N=1000*	*N=500*	*N=1000*	*N=500*	*N=1000*
Angle between Clusters = 90 ($\alpha=90^0$, $\beta=0^0$)						
All N items W/O Centering	0.44 0.03	0.38 0.02	0.87 0.10	0.82 0.09	0.97 0.11	0.97 0.11
N-2 items W/O Centering	0.48 0.04	0.48 0.03	0.82 0.12	0.78 0.11	0.90 0.14	0.91 0.12
Average W/O Centering	0.27 0.03	0.22 0.02	0.78 0.11	0.73 0.10	0.90 0.13	0.91 0.12
All N items W Centering	0.25 0.03	0.20 0.02	0.75 0.12	0.70 0.10	0.86 0.13	0.87 0.12
N-2 items W Centering	0.30 0.04	0.26 0.02	0.82 0.12	0.77 0.10	0.93 0.14	0.94 0.12
Average W Centering	0.26 0.03	0.21 0.02	0.78 0.12	0.73 0.10	0.89 0.14	0.90 0.12
Angle between Clusters = 70 ($\alpha=80^0$, $\beta=10^0$)						
All N items W/O Centering	0.43 0.03	0.37 0.02	0.63 0.07	0.60 0.05	0.68 0.06	0.66 0.06
N-2 items W/O Centering	0.46 0.04	0.47 0.02	0.52 0.07	0.50 0.04	0.54 0.07	0.53 0.07
Average W/O Centering	0.28 0.03	0.22 0.02	0.50 0.08	0.47 0.06	0.56 0.08	0.54 0.07
All N items W Centering	0.25 0.02	0.20 0.02	0.47 0.08	0.44 0.06	0.53 0.07	0.51 0.07
N-2 items W Centering	0.32 0.03	0.27 0.03	0.52 0.09	0.50 0.06	0.59 0.08	0.57 0.07
Average W Centering	0.27 0.03	0.22 0.02	0.49 0.08	0.46 0.06	0.55 0.08	0.53 0.07
Angle between Clusters = 50 ($\alpha=70^0$, $\beta=20^0$)						
All N items W/O Centering	0.42 0.02	0.37 0.02	0.48 0.04	0.45 0.04	0.50 0.03	0.46 0.03
N-2 items W/O Centering	0.46 0.03	0.47 0.02	0.46 0.03	0.46 0.02	0.45 0.03	0.46 0.02
Average W/O Centering	0.27 0.03	0.23 0.03	0.32 0.04	0.30 0.04	0.34 0.04	0.32 0.04
All N items W Centering	0.25 0.02	0.20 0.02	0.30 0.04	0.29 0.04	0.32 0.04	0.29 0.03
N-2 items W Centering	0.32 0.03	0.29 0.03	0.34 0.04	0.32 0.04	0.36 0.04	0.34 0.04
Average W Centering	0.26 0.03	0.23 0.03	0.31 0.04	0.29 0.04	0.33 0.04	0.31 0.04

(To be continued)

Table 5 (Continued): *D-max* Value for 30 Items (Standard Deviations Are Presented in Smaller Font)

Type of Conditional Score	Number of Items in Each Cluster					
	30-0		20-10		15-15	
	N=500	*N=1000*	*N=500*	*N=1000*	*N=500*	*N=1000*
Angle between Clusters = 30 ($\alpha=90^0$, $\beta=0^0$)						
All N items W/O Centering	0.42 0.03	0.36 0.03	0.42 0.03	0.36 0.03	0.43 0.02	0.37 0.02
N-2 items W/O Centering	0.45 0.04	0.46 0.02	0.46 0.04	0.46 0.02	0.45 0.03	0.46 0.02
Average W/O Centering	0.27 0.04	0.23 0.03	0.27 0.03	0.22 0.03	0.28 0.03	0.23 0.03
All N items W Centering	0.25 0.03	0.21 0.02	0.25 0.02	0.20 0.02	0.25 0.02	0.21 0.02
N-2 items W Centering	0.32 0.04	0.29 0.03	0.31 0.04	0.29 0.03	0.32 0.04	0.29 0.03
Average W Centering	0.27 0.04	0.22 0.03	0.26 0.03	0.22 0.03	0.27 0.03	0.23 0.03
Angle between Clusters = 10 ($\alpha=50^0$, $\beta=40^0$)						
All N items W/O Centering	0.41 0.03	0.36 0.03	0.42 0.03	0.36 0.02	0.42 0.03	0.36 0.03
N-2 items W/O Centering	0.45 0.03	0.46 0.03	0.45 0.03	0.46 0.02	0.46 0.03	0.45 0.02
Average W/O Centering	0.27 0.03	0.23 0.03	0.27 0.03	0.24 0.03	0.27 0.04	0.23 0.03
All N items W Centering	0.25 0.03	0.21 0.02	0.25 0.03	0.21 0.02	0.25 0.03	0.21 0.02
N-2 items W Centering	0.32 0.04	0.29 0.03	0.32 0.04	0.30 0.03	0.32 0.04	0.29 0.03
Average W Centering	0.26 0.03	0.23 0.03	0.27 0.03	0.23 0.03	0.27 0.03	0.23 0.03
Angle between Clusters = 0 ($\alpha=45^0$, $\beta=45^0$)						
All N items W/O Centering	0.41 0.03	0.36 0.03	0.42 0.03	0.36 0.03	0.41 0.03	0.36 0.02
N-2 items W/O Centering	0.45 0.03	0.46 0.02	0.46 0.03	0.46 0.02	0.45 0.04	0.46 0.02
Average W/O Centering	0.27 0.03	0.24 0.03	0.27 0.03	0.23 0.03	0.27 0.03	0.23 0.03
All N items W Centering	0.25 0.03	0.21 0.02	0.25 0.03	0.21 0.02	0.24 0.02	0.21 0.02
N-2 items W Centering	0.32 0.04	0.30 0.04	0.32 0.04	0.29 0.03	0.32 0.04	0.29 0.03
Average W Centering	0.27 0.03	0.23 0.03	0.27 0.03	0.23 0.03	0.26 0.03	0.23 0.03

Table 6: *D-max* Value for 60 Items (Standard Deviations Are Presented in Smaller Font)

Type of Conditional Score	Number of Items in Each Cluster					
	60-0		40-20		30-30	
	N=500	*N=1000*	*N=500*	*N=1000*	*N=500*	*N=1000*
Angle between Clusters = 90 ($\alpha=90^0$, $\beta=0^0$)						
All N items W/O Centering	0.27 0.01	0.22 0.01	0.49 0.04	0.49 0.03	0.71 0.05	0.72 0.05
N-2 items W/O Centering	0.25 0.04	0.25 0.02	0.49 0.05	0.50 0.03	0.70 0.05	0.71 0.06
Average W/O Centering	0.19 0.01	0.14 0.01	0.45 0.05	0.46 0.03	0.69 0.05	0.71 0.06
All N items W Centering	0.18 0.01	0.13 0.01	0.43 0.05	0.43 0.03	0.76 0.05	0.69 0.06
N-2 items W Centering	0.19 0.01	0.15 0.01	0.48 0.05	0.49 0.03	0.71 0.05	0.72 0.06
Average W Centering	0.18 0.01	0.13 0.01	0.45 0.05	0.46 0.03	0.69 0.05	0.71 0.06
Angle between Clusters = 70 ($\alpha=80^0$, $\beta=10^0$)						
All N items W/O Centering	0.26 0.01	0.21 0.01	0.36 0.03	0.34 0.03	0.45 0.04	0.45 0.04
N-2 items W/O Centering	0.24 0.03	0.24 0.02	0.32 0.03	0.31 0.03	0.40 0.04	0.41 0.05
Average W/O Centering	0.18 0.01	0.13 0.01	0.30 0.03	0.29 0.03	0.41 0.04	0.41 0.04
All N items W Centering	0.18 0.01	0.13 0.01	0.28 0.03	0.27 0.03	0.39 0.04	0.40 0.04
N-2 items W Centering	0.19 0.01	0.15 0.01	0.32 0.03	0.31 0.03	0.42 0.04	0.43 0.04
Average W Centering	0.18 0.01	0.13 0.01	0.30 0.03	0.29 0.03	0.41 0.04	0.41 0.04
Angle between Clusters = 50 ($\alpha=70^0$, $\beta=20^0$)						
All N items W/O Centering	0.26 0.01	0.21 0.01	0.28 0.02	0.25 0.02	0.32 0.02	0.29 0.02
N-2 items W/O Centering	0.23 0.03	0.24 0.02	0.24 0.03	0.24 0.02	0.25 0.02	0.24 0.01
Average W/O Centering	0.18 0.01	0.13 0.01	0.21 0.02	0.19 0.02	0.24 0.03	0.23 0.02
All N items W Centering	0.17 0.01	0.13 0.01	0.20 0.02	0.17 0.02	0.24 0.02	0.22 0.02
N-2 items W Centering	0.19 0.01	0.15 0.01	0.22 0.02	0.20 0.02	0.26 0.03	0.24 0.03
Average W Centering	0.18 0.01	0.13 0.01	0.21 0.02	0.18 0.02	0.24 0.02	0.22 0.02

(To be continued)

Table 6 (Continued): *D-max* Value for 60 Items (Standard Deviations Are Presented in Smaller Font)

Type of Conditional Score	Number of Items in Each Cluster					
	60-0		40-20		30-30	
	N=500	*N=1000*	*N=500*	*N=1000*	*N=500*	*N=1000*
Angle between Clusters = 30 ($\alpha=90^0$, $\beta=0^0$)						
All N items W/O Centering	0.25 0.01	0.21 0.01	0.25 0.01	0.21 0.01	0.26 0.01	0.22 0.01
N-2 items W/O Centering	0.23 0.03	0.24 0.02	0.24 0.02	0.23 0.02	0.23 0.03	0.23 0.02
Average W/O Centering	0.18 0.01	0.14 0.01	0.18 0.02	0.14 0.01	0.18 0.01	0.14 0.01
All N items W Centering	0.17 0.01	0.13 0.01	0.17 0.01	0.13 0.01	0.18 0.01	0.14 0.01
N-2 items W Centering	0.19 0.02	0.16 0.01	0.19 0.02	0.16 0.01	0.19 0.01	0.15 0.01
Average W Centering	0.17 0.01	0.13 0.01	0.17 0.02	0.14 0.01	0.18 0.01	0.14 0.01
Angle between Clusters = 10 ($\alpha=50^0$, $\beta=40^0$)						
All N items W/O Centering	0.25 0.01	0.21 0.01	0.25 0.01	0.21 0.01	0.25 0.01	0.21 0.01
N-2 items W/O Centering	0.23 0.03	0.23 0.02	0.23 0.03	0.23 0.02	0.23 0.03	0.23 0.02
Average W/O Centering	0.17 0.01	0.14 0.01	0.17 0.01	0.14 0.01	0.18 0.01	0.14 0.01
All N items W Centering	0.17 0.01	0.13 0.01	0.17 0.01	0.13 0.01	0.17 0.01	0.13 0.01
N-2 items W Centering	0.19 0.01	0.16 0.01	0.19 0.01	0.16 0.01	0.19 0.01	0.16 0.01
Average W Centering	0.17 0.01	0.13 0.01	0.17 0.01	0.13 0.01	0.17 0.01	0.13 0.01
Angle between Clusters = 0 ($\alpha=45^0$, $\beta=45^0$)						
All N items W/O Centering	0.25 0.01	0.21 0.01	0.25 0.01	0.21 0.01	0.25 0.01	0.21 0.01
N-2 items W/O Centering	0.23 0.03	0.23 0.02	0.23 0.02	0.23 0.02	0.23 0.03	0.23 0.02
Average W/O Centering	0.17 0.01	0.14 0.01	0.17 0.01	0.14 0.01	0.18 0.01	0.14 0.01
All N items W Centering	0.17 0.01	0.13 0.01	0.17 0.01	0.13 0.01	0.17 0.01	0.13 0.01
N-2 items W Centering	0.19 0.01	0.16 0.01	0.19 0.01	0.16 0.01	0.19 0.01	0.16 0.01
Average W Centering	0.17 0.01	0.13 0.01	0.17 0.01	0.13 0.01	0.17 0.01	0.13 0.01

Table 7: *r* Value for 30 Items (Standard Deviations Are Presented in Smaller Font)

Type of Conditional Score	Number of Items in Each Cluster					
	30-0		20-10		15-15	
	N=500	*N=1000*	*N=500*	*N=1000*	*N=500*	*N=1000*
Angle between Clusters = 90 ($\alpha=90^0$, $\beta=0^0$)						
All N items W/O Centering	0.59 0.02	0.58 0.02	0.80 0.04	0.84 0.03	0.84 0.04	0.89 0.03
N-2 items W/O Centering	0.61 0.05	0.74 0.02	0.77 0.06	0.84 0.05	0.79 0.06	0.88 0.04
Average W/O Centering	0.43 0.04	0.46 0.04	0.79 0.05	0.86 0.04	0.84 0.05	0.92 0.03
All N items W Centering	0.40 0.03	0.42 0.04	0.76 0.06	0.83 0.05	0.82 0.05	0.90 0.03
N-2 items W Centering	0.43 0.04	0.49 0.04	0.79 0.05	0.87 0.03	0.84 0.05	0.92 0.03
Average W Centering	0.41 0.04	0.45 0.04	0.78 0.05	0.86 0.04	0.83 0.05	0.92 0.03
Angle between Clusters = 70 ($\alpha=80^0$, $\beta=10^0$)						
All N items W/O Centering	0.59 0.03	0.58 0.02	0.71 0.04	0.76 0.03	0.74 0.04	0.80 0.03
N-2 items W/O Centering	0.60 0.03	0.74 0.02	0.60 0.05	0.68 0.03	0.61 0.05	0.69 0.05
Average W/O Centering	0.44 0.04	0.48 0.04	0.65 0.06	0.75 0.04	0.69 0.05	0.80 0.04
All N items W Centering	0.41 0.03	0.43 0.04	0.62 0.06	0.70 0.05	0.66 0.06	0.75 0.05
N-2 items W Centering	0.46 0.04	0.53 0.04	0.64 0.07	0.74 0.05	0.69 0.06	0.79 0.05
Average W Centering	0.43 0.04	0.47 0.04	0.64 0.06	0.74 0.05	0.68 0.06	0.79 0.05
Angle between Clusters = 50 ($\alpha=70^0$, $\beta=20^0$)						
All N items W/O Centering	0.59 0.02	0.59 0.02	0.63 0.03	0.66 0.04	0.64 0.03	0.68 0.03
N-2 items W/O Centering	0.61 0.03	0.74 0.02	0.59 0.04	0.70 0.03	0.58 0.04	0.70 0.03
Average W/O Centering	0.45 0.04	0.50 0.05	0.50 0.05	0.59 0.06	0.53 0.05	0.61 0.05
All N items W Centering	0.42 0.03	0.44 0.04	0.47 0.05	0.55 0.06	0.49 0.05	0.55 0.05
N-2 items W Centering	0.48 0.04	0.55 0.04	0.48 0.05	0.56 0.06	0.51 0.05	0.58 0.05
Average W Centering	0.43 0.04	0.49 0.04	0.48 0.05	0.57 0.06	0.51 0.05	0.59 0.05

(To be continued)

Table 7 (Continued): *r* Value for 30 Items (Standard Deviations Are Presented in Smaller Font)

Type of Conditional Score	Number of Items in Each Cluster					
	30-0		20-10		15-15	
	N=500	*N=1000*	*N=500*	*N=1000*	*N=500*	*N=1000*
Angle between Clusters = 30 ($\alpha=90^0$, $\beta=0^0$)						
All N items W/O Centering	0.59	0.59	0.59	0.59	0.59	0.59
	0.03	0.03	0.03	0.03	0.02	0.03
N-2 items W/O Centering	0.61	0.73	0.61	0.73	0.60	0.73
	0.04	0.03	0.03	0.02	0.03	0.03
Average W/O Centering	0.46	0.51	0.45	0.49	0.45	0.50
	0.04	0.05	0.03	0.05	0.04	0.05
All N items W Centering	0.42	0.45	0.41	0.44	0.42	0.44
	0.04	0.04	0.03	0.04	0.03	0.04
N-2 items W Centering	0.49	0.55	0.47	0.55	0.48	0.55
	0.04	0.04	0.05	0.05	0.04	0.04
Average W Centering	0.44	0.50	0.43	0.49	0.44	0.49
	0.04	0.05	0.04	0.05	0.04	0.05
Angle between Clusters = 10 ($\alpha=50^0$, $\beta=40^0$)						
All N items W/O Centering	0.59	0.59	0.59	0.59	0.59	0.59
	0.02	0.03	0.02	0.03	0.03	0.02
N-2 items W/O Centering	0.61	0.73	0.61	0.73	0.61	0.73
	0.04	0.03	0.03	0.03	0.03	0.03
Average W/O Centering	0.45	0.52	0.46	0.52	0.46	0.51
	0.04	0.05	0.04	0.05	0.05	0.05
All N items W Centering	0.42	0.46	0.42	0.46	0.42	0.46
	0.04	0.04	0.04	0.04	0.04	0.04
N-2 items W Centering	0.49	0.56	0.49	0.57	0.49	0.56
	0.04	0.05	0.05	0.04	0.05	0.04
Average W Centering	0.45	0.51	0.45	0.51	0.44	0.50
	0.04	0.05	0.04	0.05	0.05	0.05
Angle between Clusters = 0 ($\alpha=45^0$, $\beta=45^0$)						
All N items W/O Centering	0.59	0.59	0.59	0.59	0.59	0.59
	0.03	0.03	0.03	0.02	0.02	0.03
N-2 items W/O Centering	0.61	0.73	0.61	0.73	0.61	0.73
	0.03	0.03	0.03	0.03	0.04	0.03
Average W/O Centering	0.46	0.52	0.46	0.51	0.46	0.52
	0.04	0.05	0.04	0.04	0.04	0.05
All N items W Centering	0.43	0.46	0.42	0.45	0.42	0.46
	0.04	0.04	0.04	0.04	0.03	0.04
N-2 items W Centering	0.49	0.56	0.49	0.56	0.48	0.56
	0.05	0.05	0.05	0.04	0.04	0.04
Average W Centering	0.45	0.51	0.45	0.50	0.44	0.51
	0.05	0.05	0.04	0.05	0.04	0.05

Table 8: *r* Value for 60 Items (Standard Deviations Are Presented in Smaller Font)

Type of Conditional Score	Number of Items in Each Cluster					
	60-0		40-20		30-30	
	N=500	*N=1000*	*N=500*	*N=1000*	*N=500*	*N=1000*
Angle between Clusters = 90 ($\alpha=90^0$, $\beta=0^0$)						
All N items W/O Centering	0.43	0.46	0.64	0.74	0.77	0.86
	0.01	0.01	0.03	0.02	0.03	0.02
N-2 items W/O Centering	0.38	0.50	0.62	0.75	0.75	0.86
	0.05	0.04	0.05	0.03	0.03	0.03
Average W/O Centering	0.31	0.31	0.61	0.73	0.77	0.87
	0.02	0.02	0.05	0.03	0.03	0.02
All N items W Centering	0.30	0.30	0.58	0.70	0.76	0.87
	0.02	0.02	0.05	0.03	0.03	0.03
N-2 items W Centering	0.31	0.32	0.62	0.75	0.77	0.87
	0.02	0.02	0.05	0.03	0.03	0.02
Average W Centering	0.31	0.31	0.61	0.73	0.77	0.87
	0.02	0.02	0.05	0.03	0.03	0.02
Angle between Clusters = 70 ($\alpha=80^0$, $\beta=10^0$)						
All N items W/O Centering	0.44	0.46	0.54	0.64	0.63	0.73
	0.02	0.01	0.03	0.03	0.03	0.03
N-2 items W/O Centering	0.38	0.50	0.46	0.57	0.55	0.67
	0.04	0.03	0.04	0.04	0.04	0.05
Average W/O Centering	0.31	0.32	0.47	0.58	0.59	0.71
	0.02	0.02	0.03	0.04	0.04	0.04
All N items W Centering	0.30	0.31	0.45	0.55	0.57	0.69
	0.02	0.02	0.04	0.04	0.04	0.04
N-2 items W Centering	0.32	0.33	0.49	0.60	0.59	0.71
	0.02	0.02	0.03	0.03	0.04	0.04
Average W Centering	0.31	0.31	0.47	0.58	0.58	0.71
	0.02	0.02	0.03	0.04	0.04	0.04
Angle between Clusters = 50 ($\alpha=70^0$, $\beta=20^0$)						
All N items W/O Centering	0.43	0.46	0.46	0.53	0.50	0.59
	0.02	0.01	0.02	0.02	0.03	0.03
N-2 items W/O Centering	0.38	0.51	0.38	0.48	0.38	0.48
	0.04	0.03	0.04	0.03	0.03	0.02
Average W/O Centering	0.32	0.33	0.36	0.43	0.41	0.50
	0.02	0.02	0.03	0.03	0.04	0.04
All N items W Centering	0.31	0.32	0.34	0.40	0.39	0.47
	0.02	0.02	0.03	0.03	0.03	0.04
N-2 items W Centering	0.32	0.35	0.37	0.44	0.41	0.49
	0.02	0.02	0.03	0.03	0.04	0.04
Average W Centering	0.31	0.32	0.36	0.42	0.40	0.49
	0.02	0.02	0.03	0.03	0.03	0.04

(To be continued)

Table 8 (Continued): *r* Value for 60 Items (Standard Deviations Are Presented in Smaller Font)

Type of Conditional Score	Number of Items in Each Cluster					
	60-0		40-20		30-30	
	N=500	*N=1000*	*N=500*	*N=1000*	*N=500*	*N=1000*
Angle between Clusters = 30 (α=90^0, β=0^0)						
All N items W/O Centering	0.44 0.02	0.46 0.02	0.44 0.02	0.47 0.02	0.44 0.02	0.48 0.02
N-2 items W/O Centering	0.39 0.04	0.51 0.03	0.39 0.04	0.50 0.03	0.38 0.04	0.50 0.03
Average W/O Centering	0.32 0.02	0.33 0.02	0.32 0.02	0.35 0.03	0.32 0.02	0.34 0.03
All N items W Centering	0.31 0.02	0.32 0.02	0.31 0.02	0.33 0.03	0.31 0.02	0.33 0.02
N-2 items W Centering	0.33 0.02	0.36 0.02	0.33 0.02	0.37 0.02	0.33 0.02	0.35 0.03
Average W Centering	0.31 0.02	0.33 0.02	0.32 0.02	0.34 0.03	0.32 0.02	0.33 0.03
Angle between Clusters = 10 (α=50^0, β=40^0)						
All N items W/O Centering	0.44 0.02	0.47 0.02	0.44 0.02	0.46 0.02	0.44 0.01	0.47 0.02
N-2 items W/O Centering	0.39 0.04	0.50 0.03	0.39 0.05	0.51 0.03	0.39 0.04	0.51 0.03
Average W/O Centering	0.32 0.02	0.34 0.03	0.32 0.02	0.34 0.02	0.32 0.02	0.34 0.02
All N items W Centering	0.31 0.02	0.33 0.03	0.31 0.02	0.33 0.02	0.31 0.02	0.33 0.02
N-2 items W Centering	0.33 0.02	0.37 0.02	0.33 0.02	0.37 0.02	0.33 0.02	0.37 0.03
Average W Centering	0.32 0.02	0.33 0.02	0.32 0.02	0.33 0.02	0.32 0.02	0.34 0.03
Angle between Clusters = 0 (α=45^0, β=45^0)						
All N items W/O Centering	0.44 0.01	0.47 0.02	0.44 0.02	0.46 0.02	0.44 0.02	0.47 0.02
N-2 items W/O Centering	0.39 0.04	0.51 0.03	0.39 0.03	0.50 0.03	0.39 0.04	0.50 0.03
Average W/O Centering	0.32 0.02	0.34 0.02	0.32 0.02	0.34 0.02	0.32 0.02	0.34 0.03
All N items W Centering	0.31 0.02	0.33 0.02	0.31 0.02	0.32 0.02	0.31 0.02	0.33 0.02
N-2 items W Centering	0.33 0.02	0.36 0.03	0.33 0.02	0.36 0.02	0.33 0.02	0.37 0.02
Average W Centering	0.31 0.02	0.33 0.02	0.32 0.02	0.33 0.02	0.32 0.02	0.34 0.03

Figure 1: Description of Item Angle

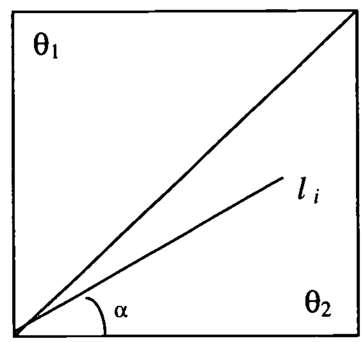


Figure 2: Example of a Test with Two Item Clusters

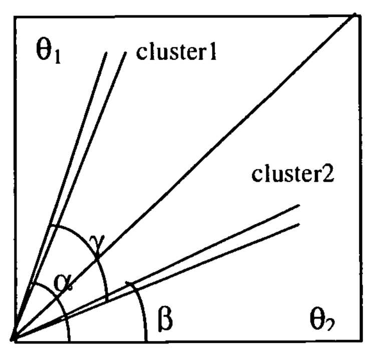


Figure 3: Percentage of Items Correctly Classified into Clusters -- 30 items

Figure 3a:

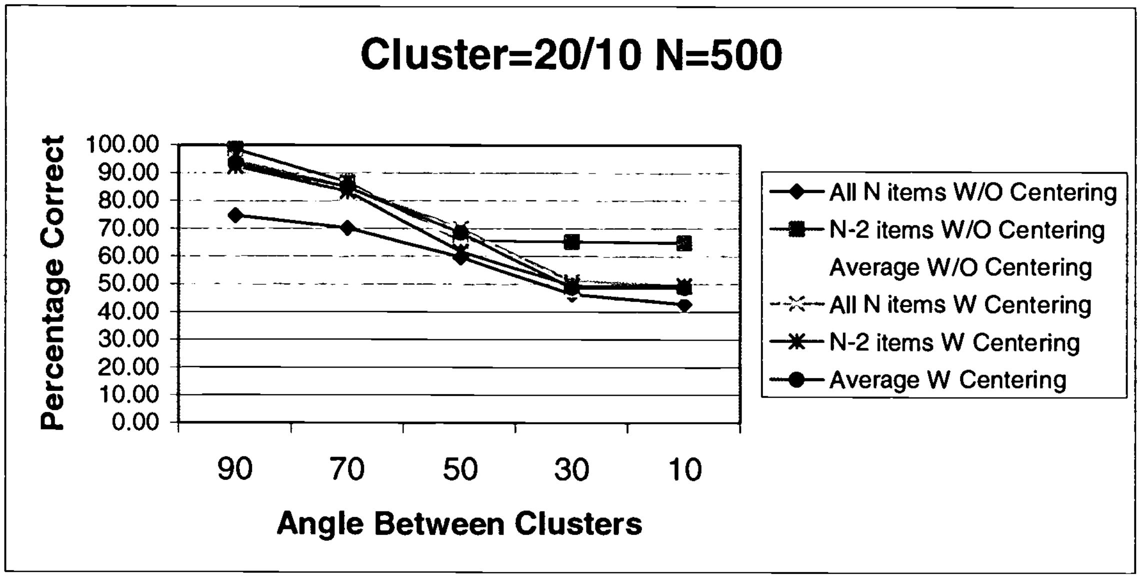


Figure 3b:

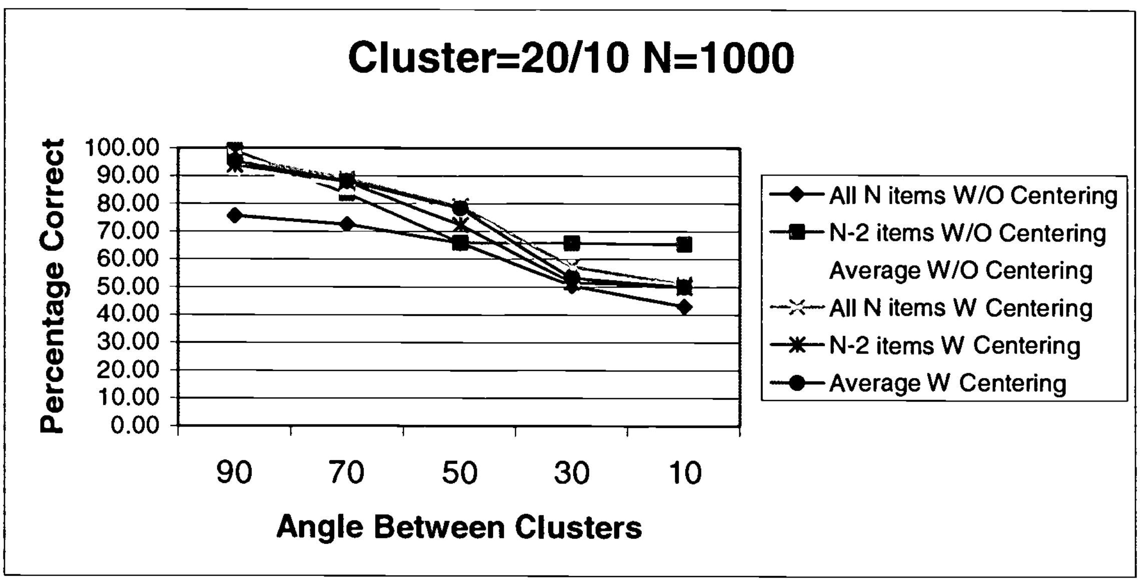


Figure 3c:

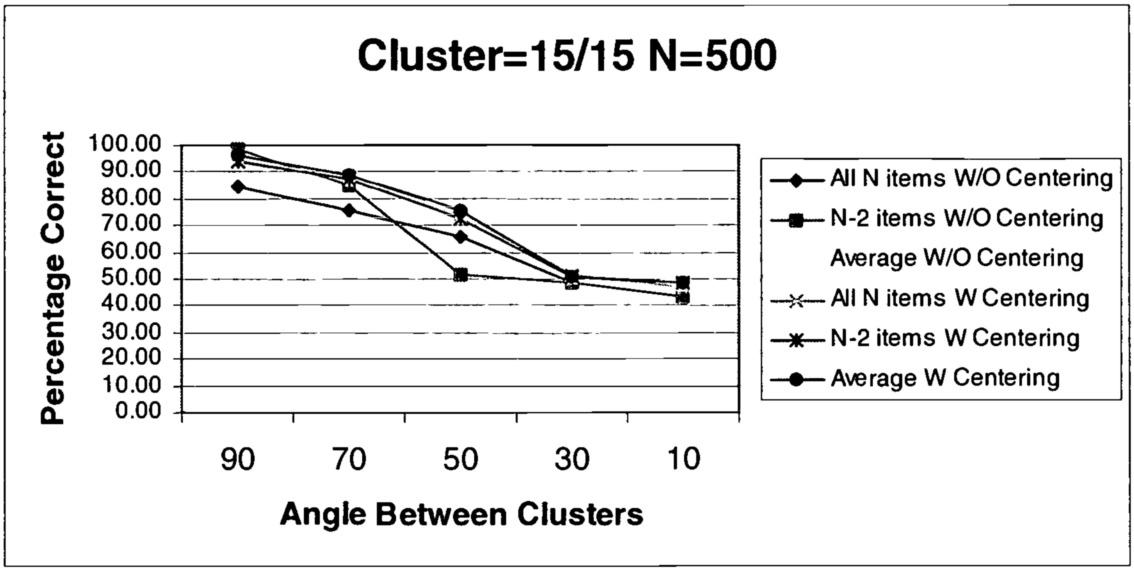
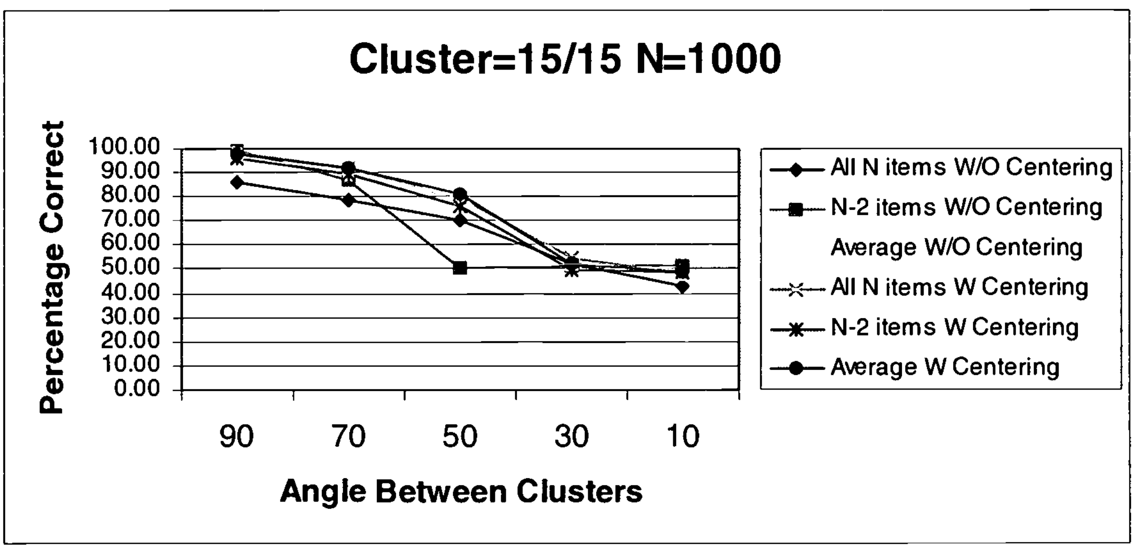


Figure 3d:



36

Figure 4. Percentage of Items Correctly Classified into Clusters -- 60 items

Figure 4a:

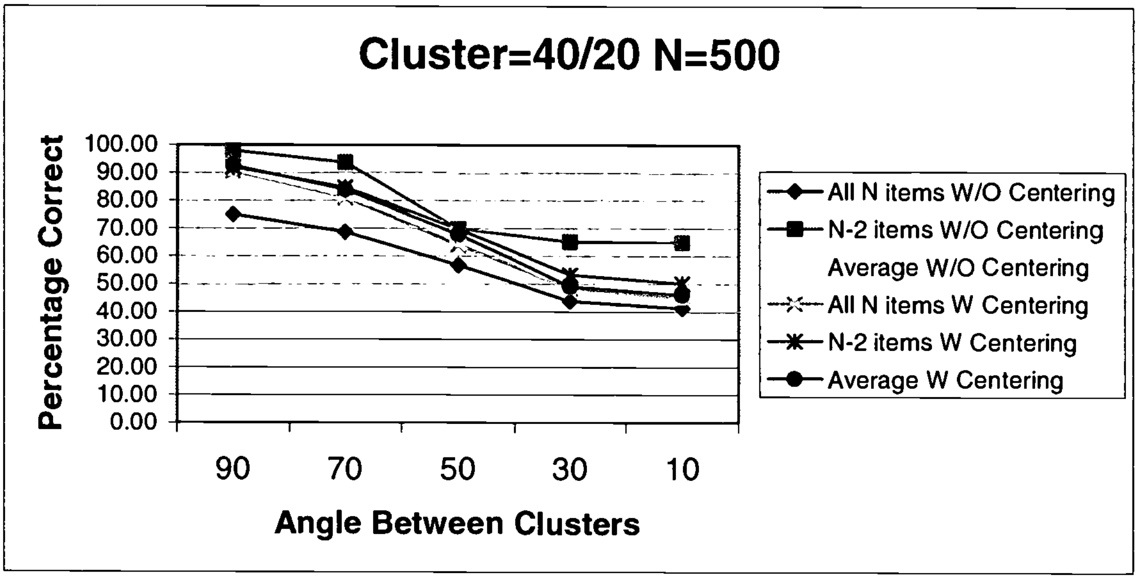


Figure 4b:

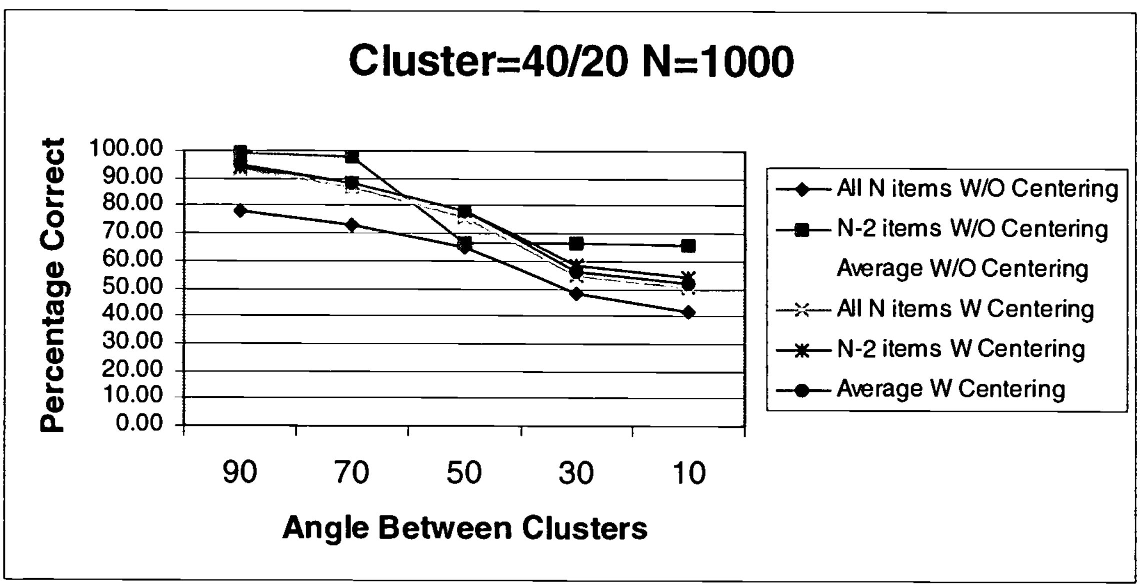


Figure 4c:

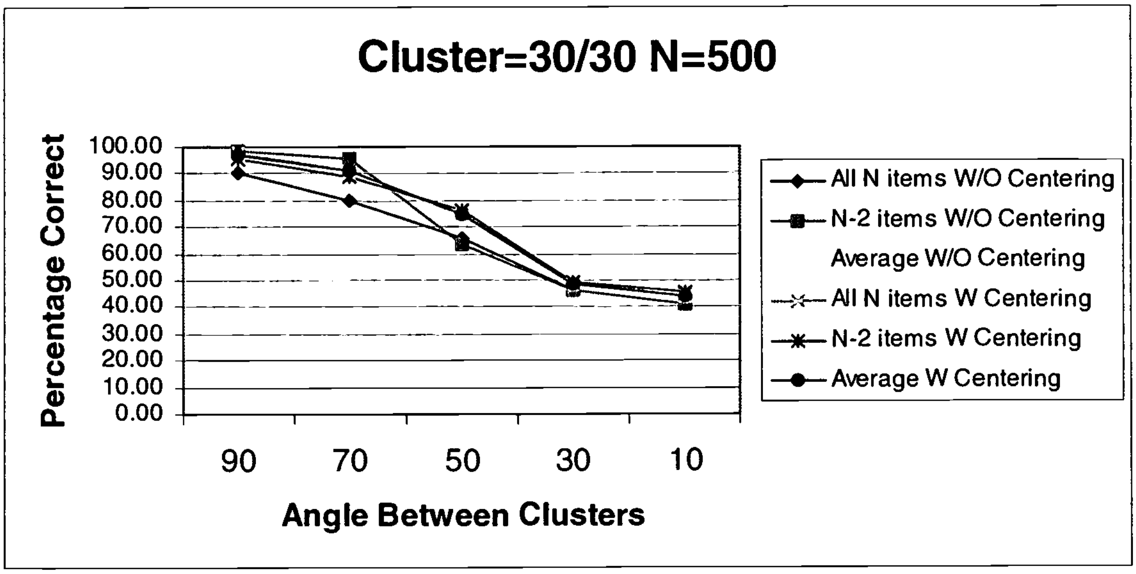


Figure 4d:

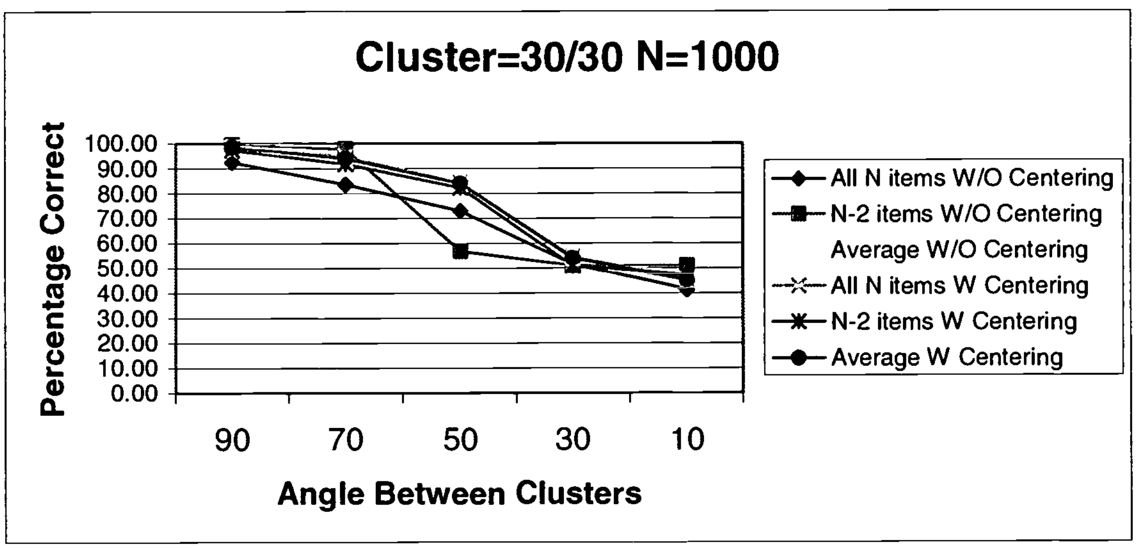


Figure 5: *D-max* Values for 30 Items

Figure 5a:

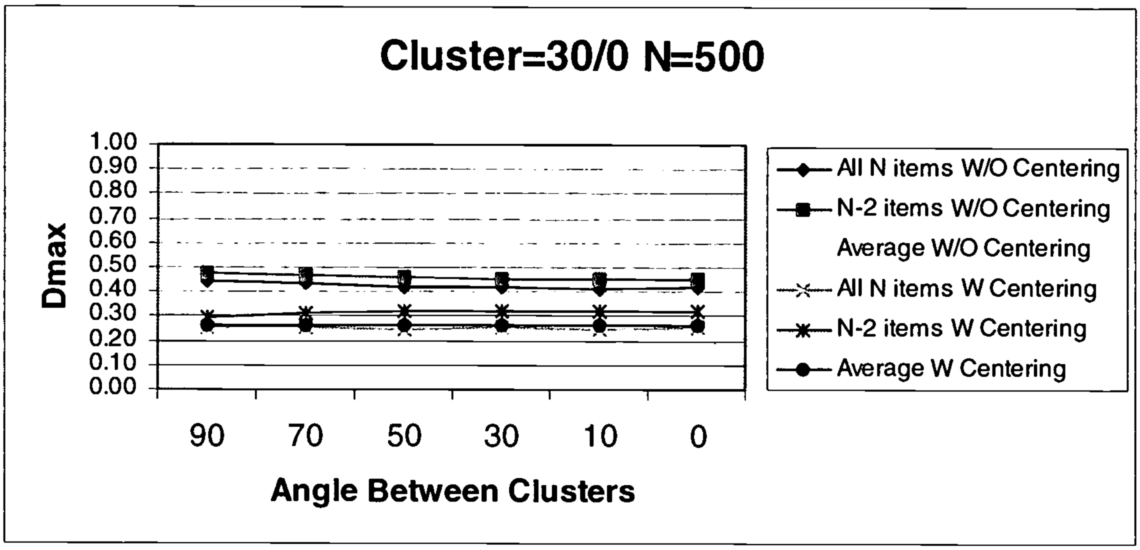


Figure 5b:

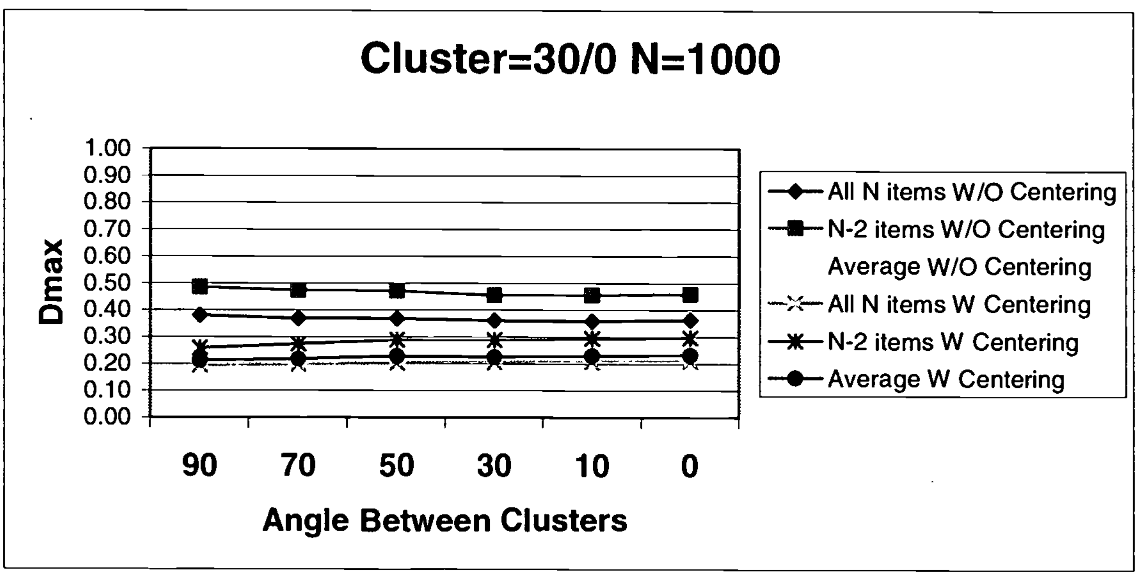


Figure 5c:

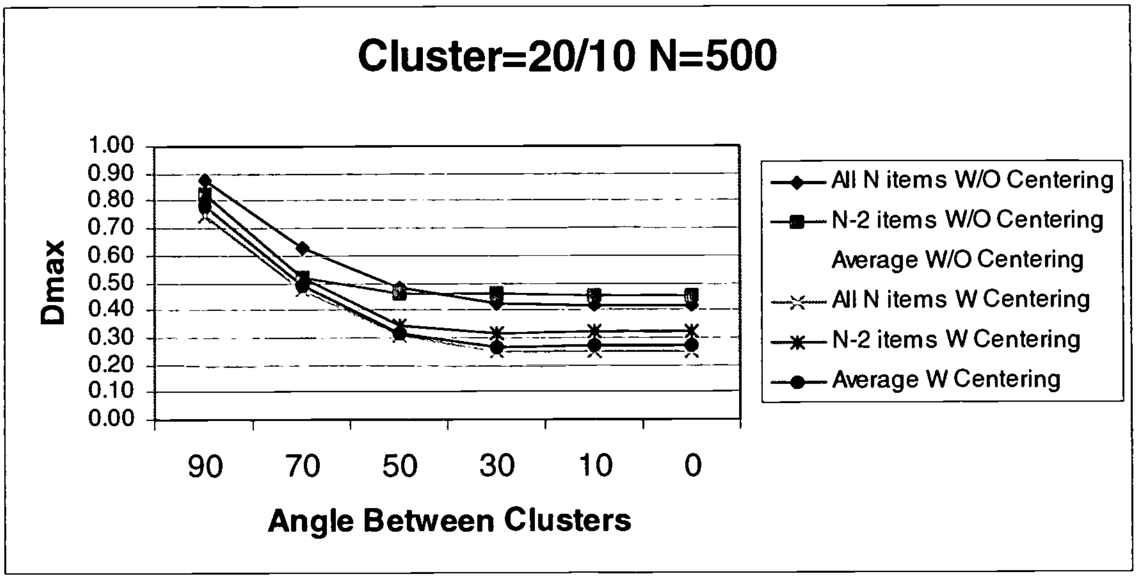


Figure 5d:

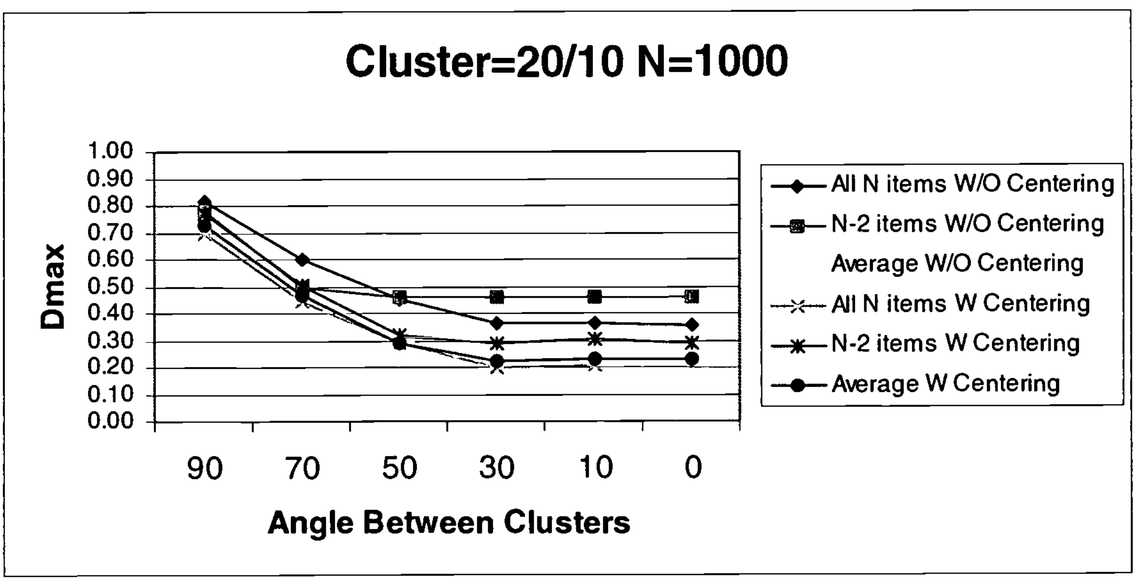


Figure 5e:

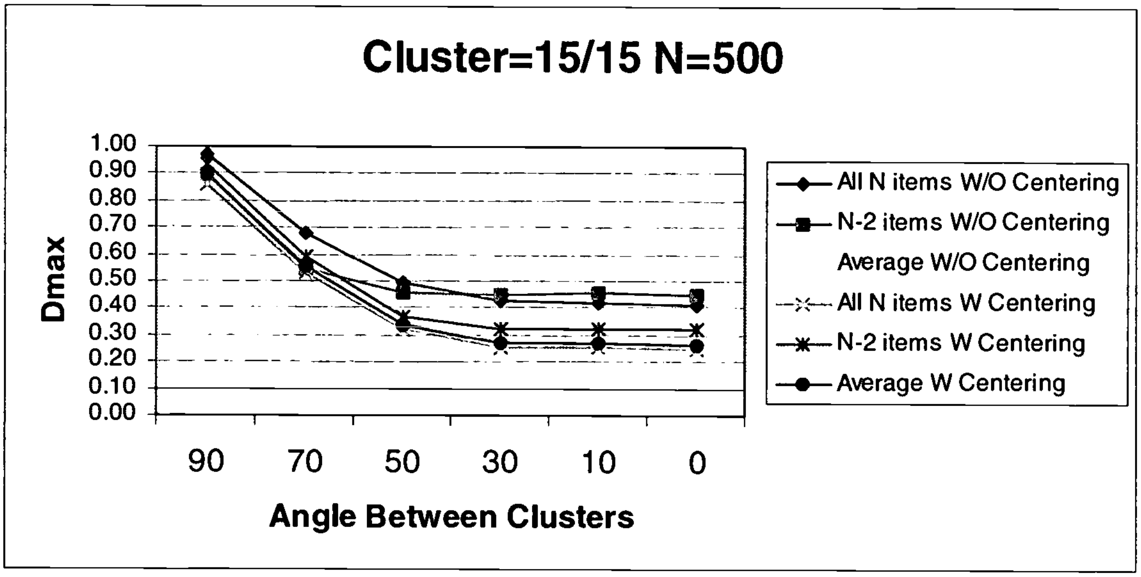


Figure 5f:

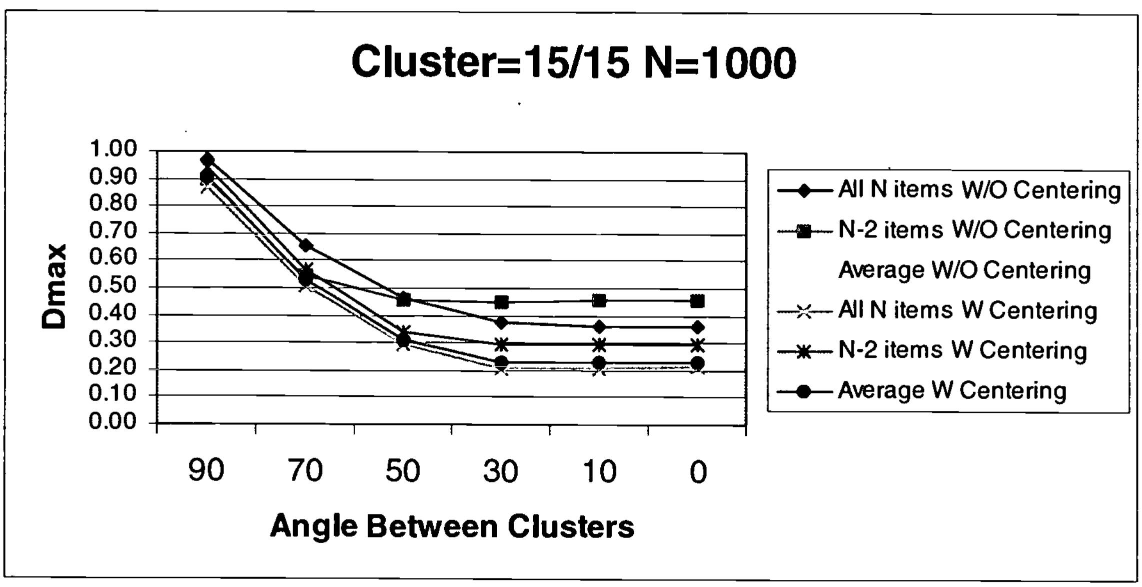


Figure 6: *D-max* Values for 60 Items

Figure 6a:

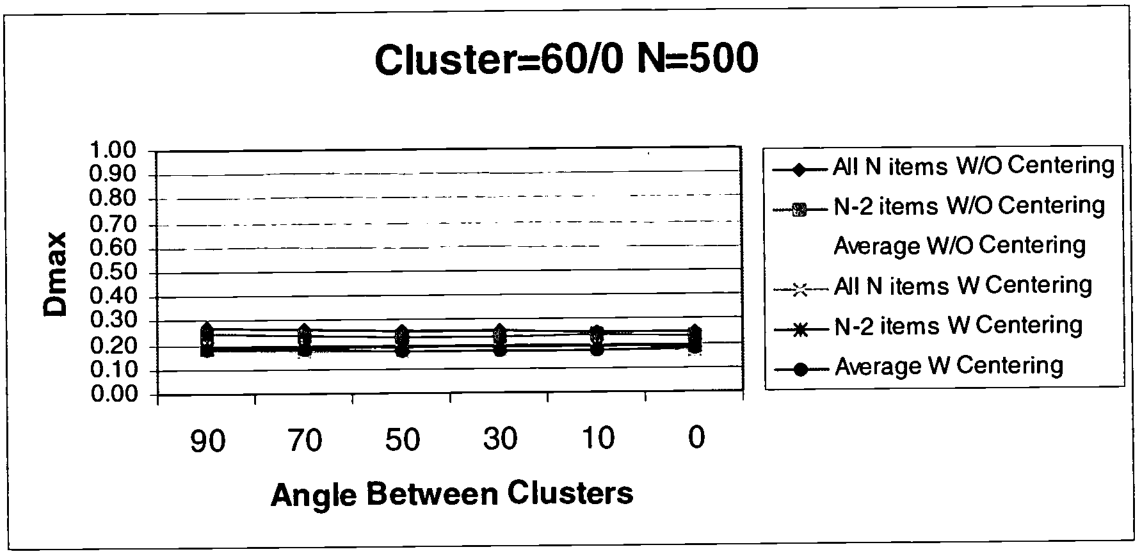
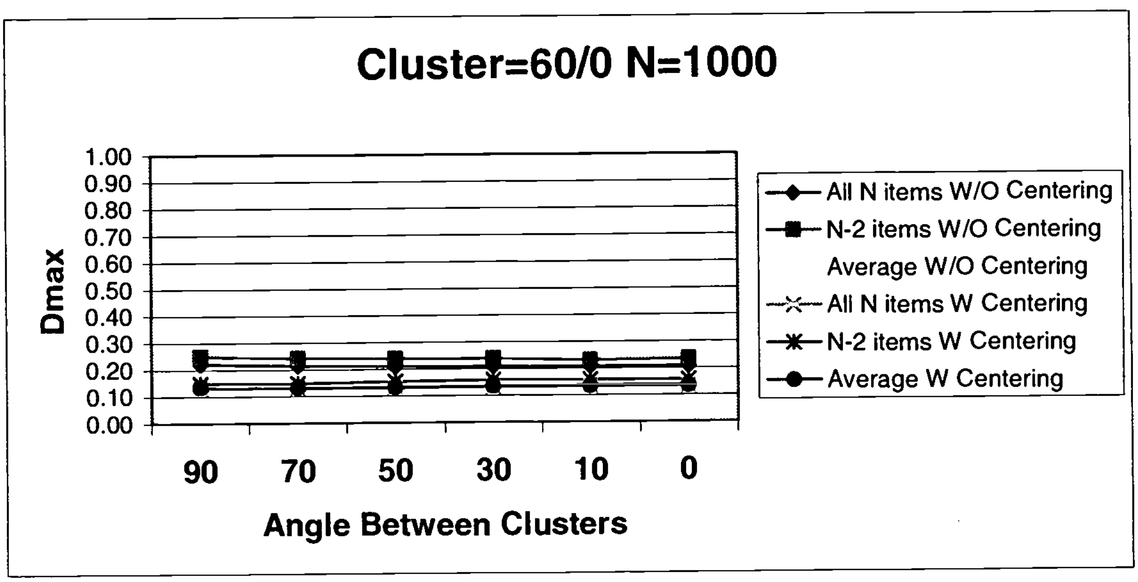


Figure 6b:



42

Figure 6c:

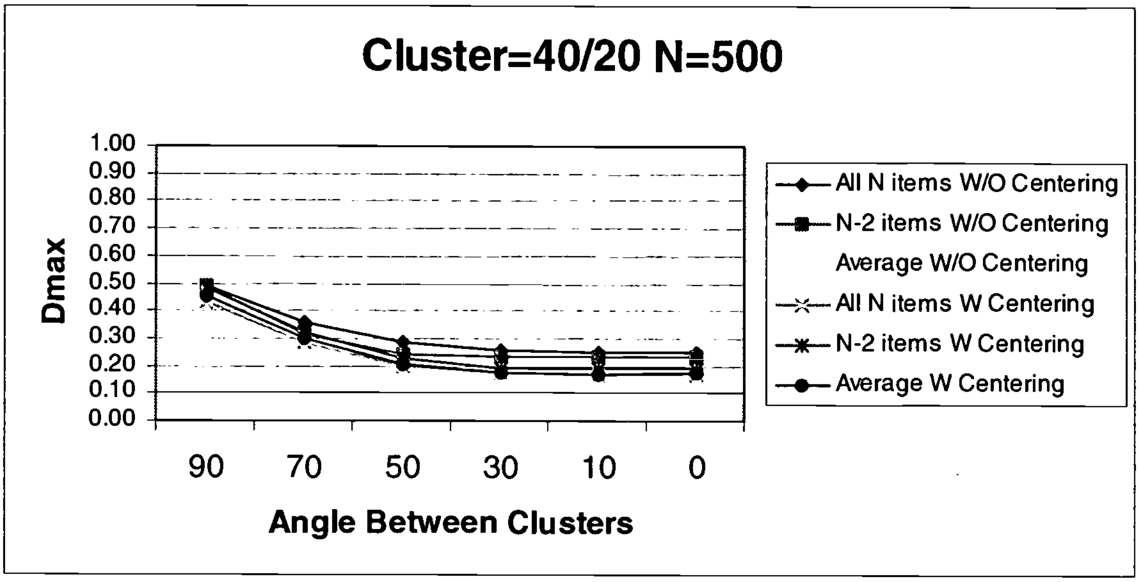


Figure 6d:

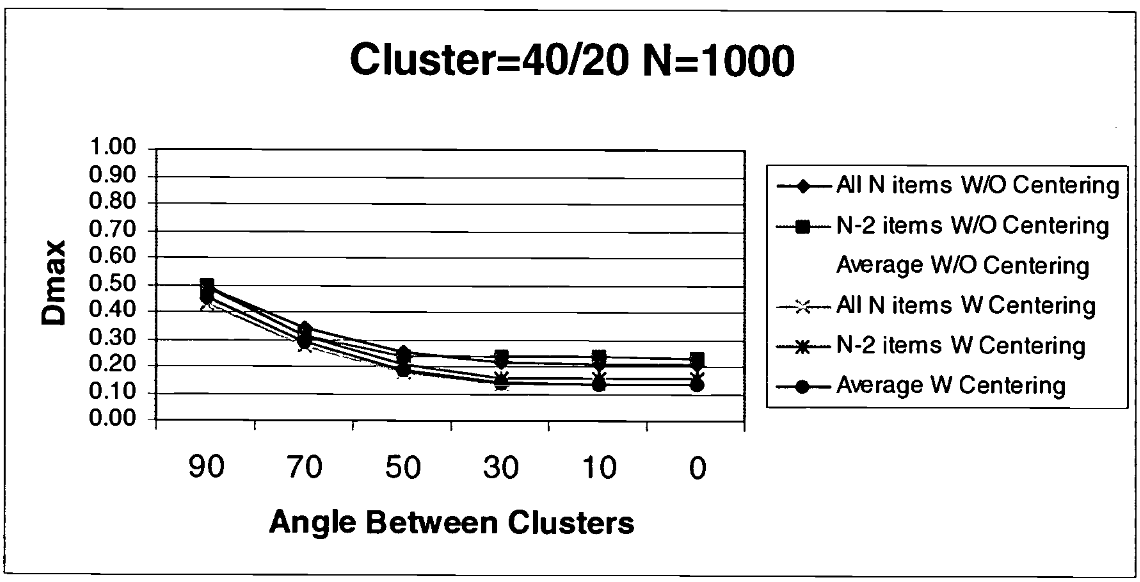


Figure 6e:

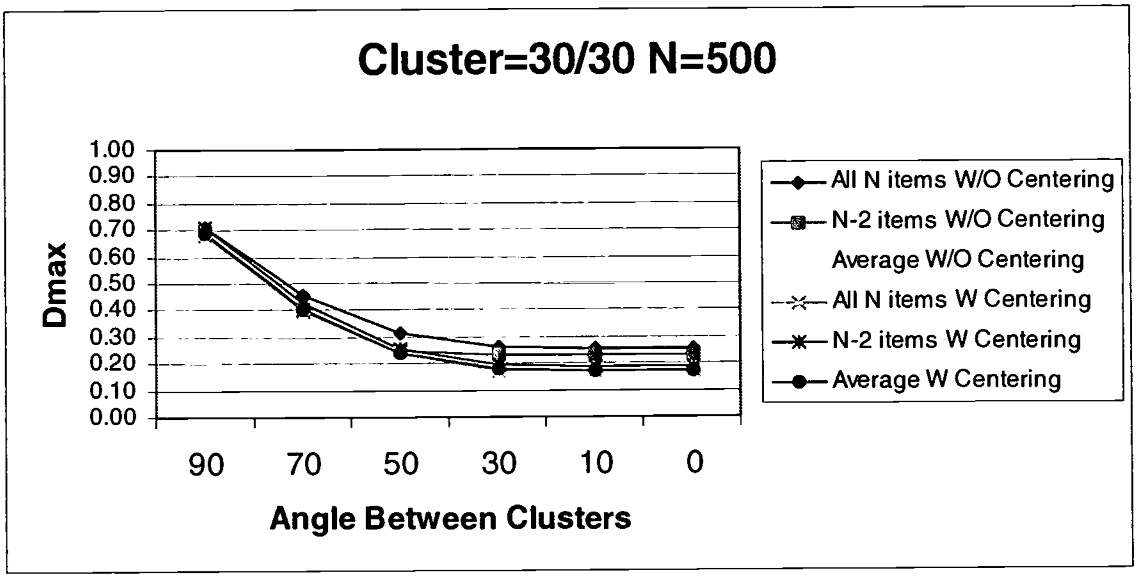


Figure 6f:

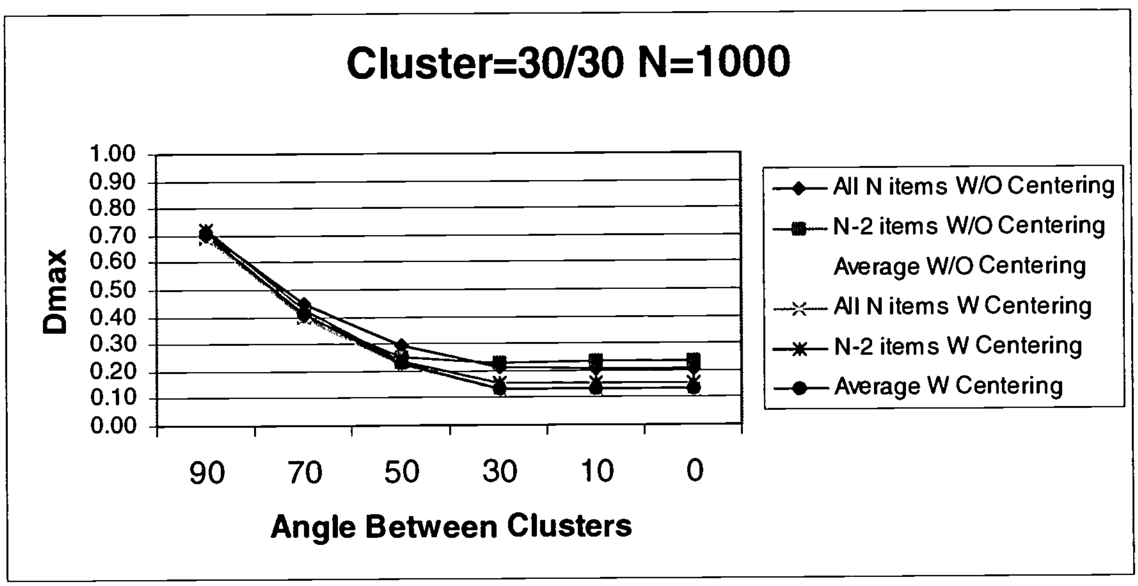


Figure 7: R Ratio for 30 Items

Figure 7a:

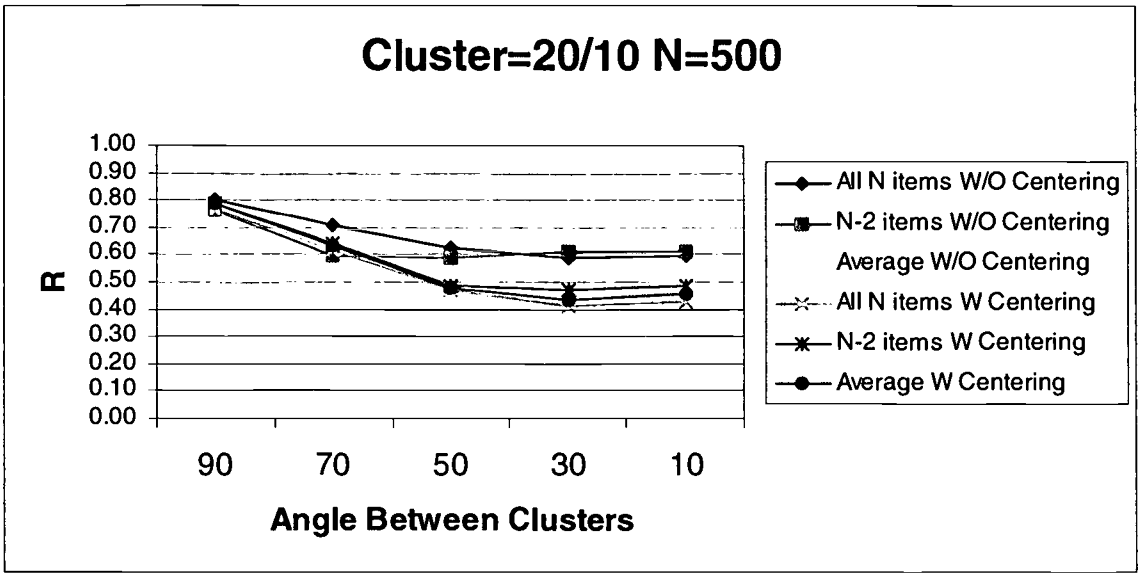


Figure 7b:

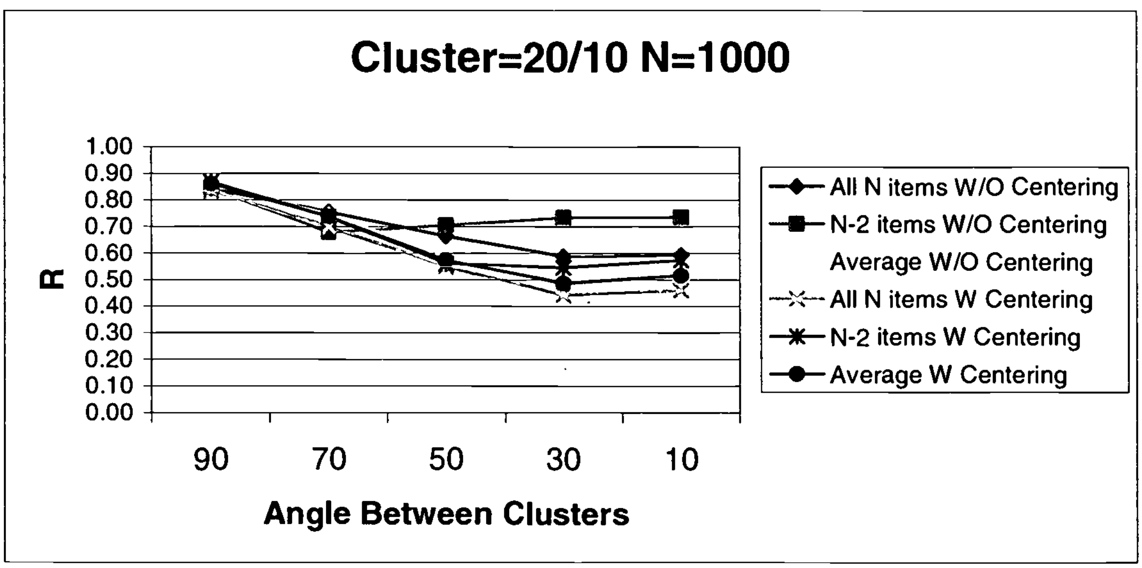


Figure 7c:

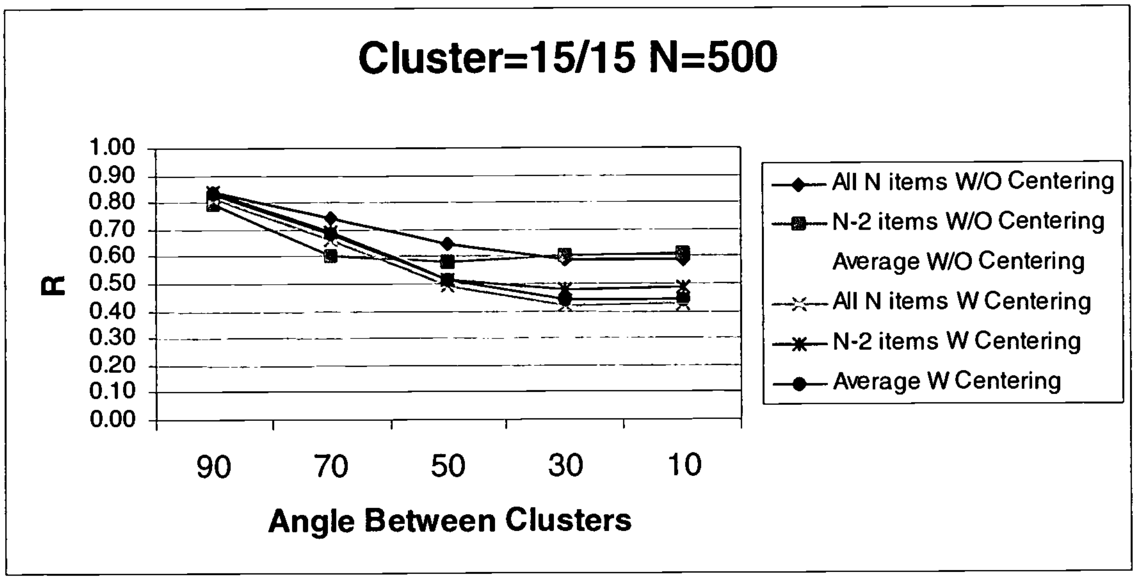


Figure 7d:

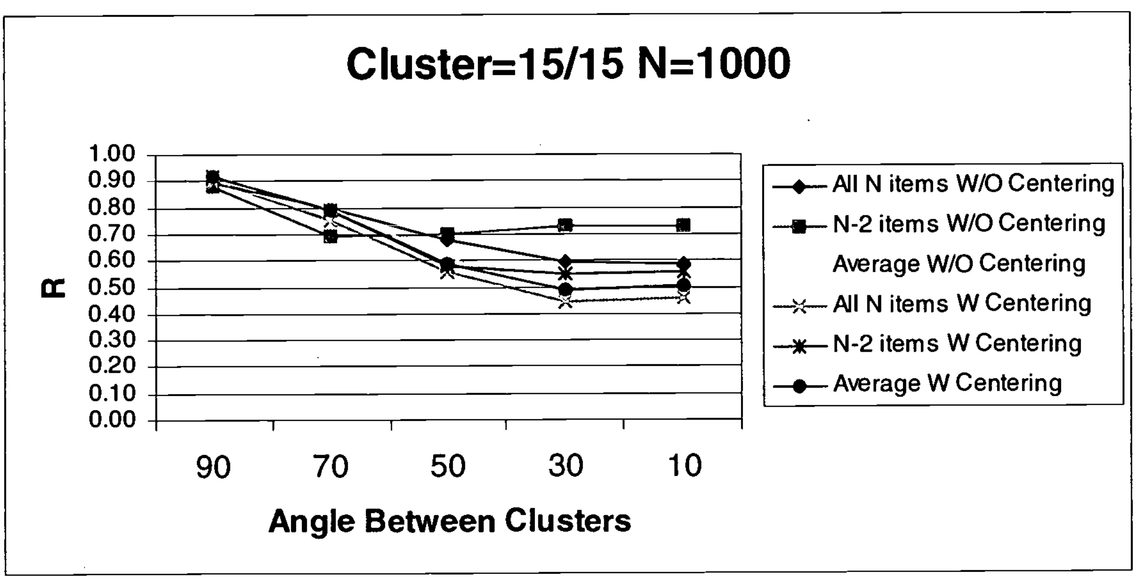


Figure 8: R Ratio for 60 Items

Figure 8a:

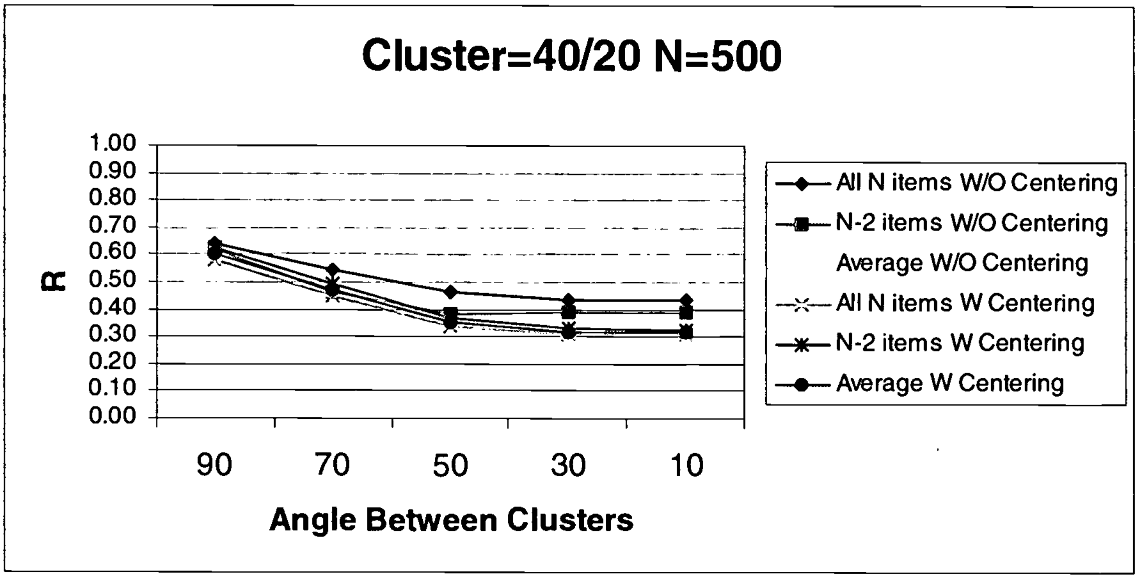


Figure 8b:

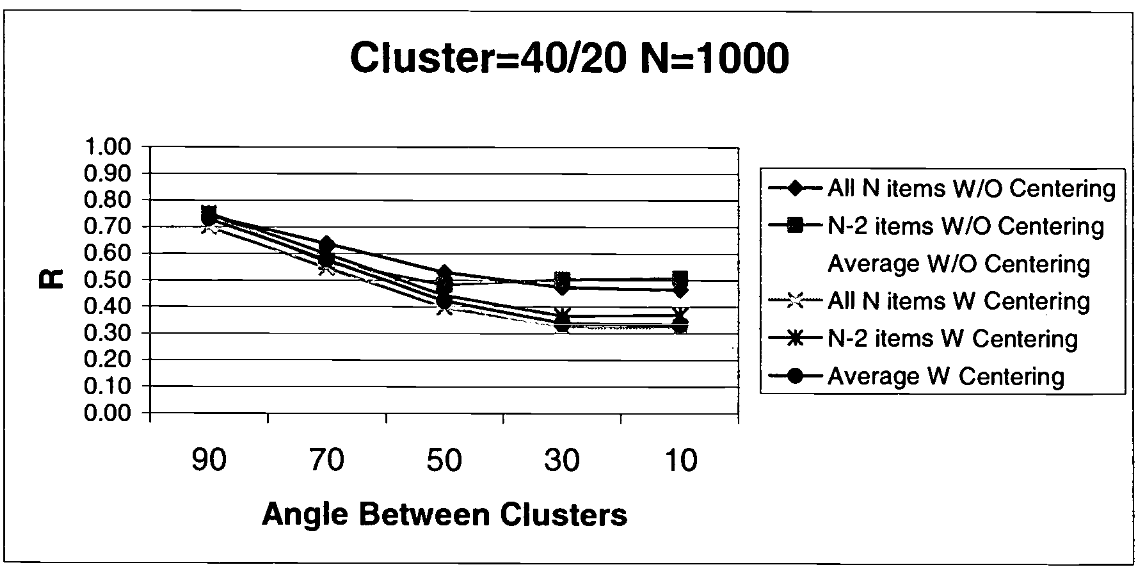


Figure 8c:

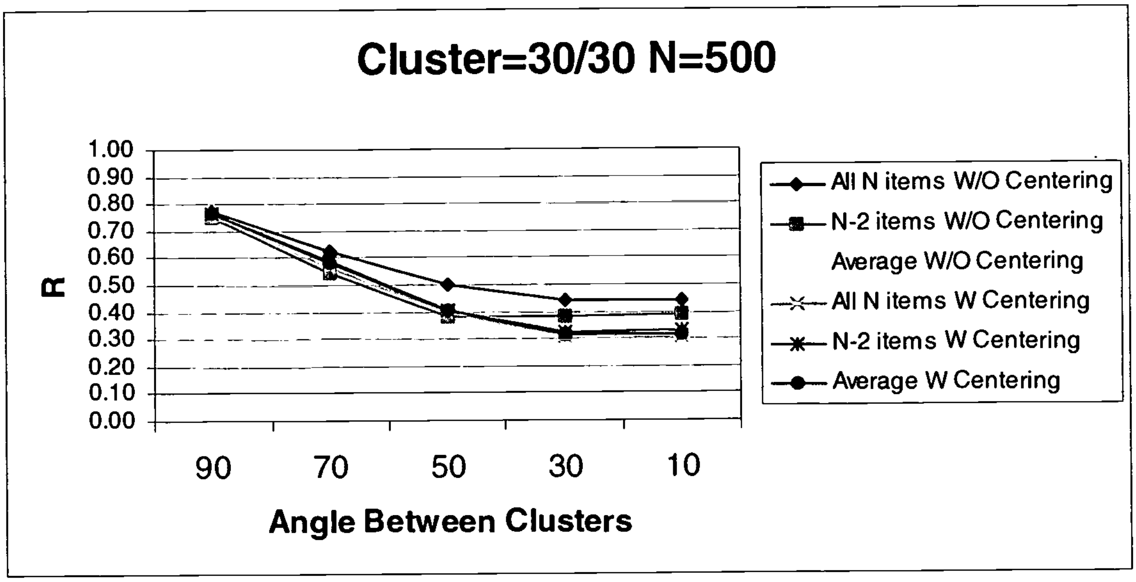
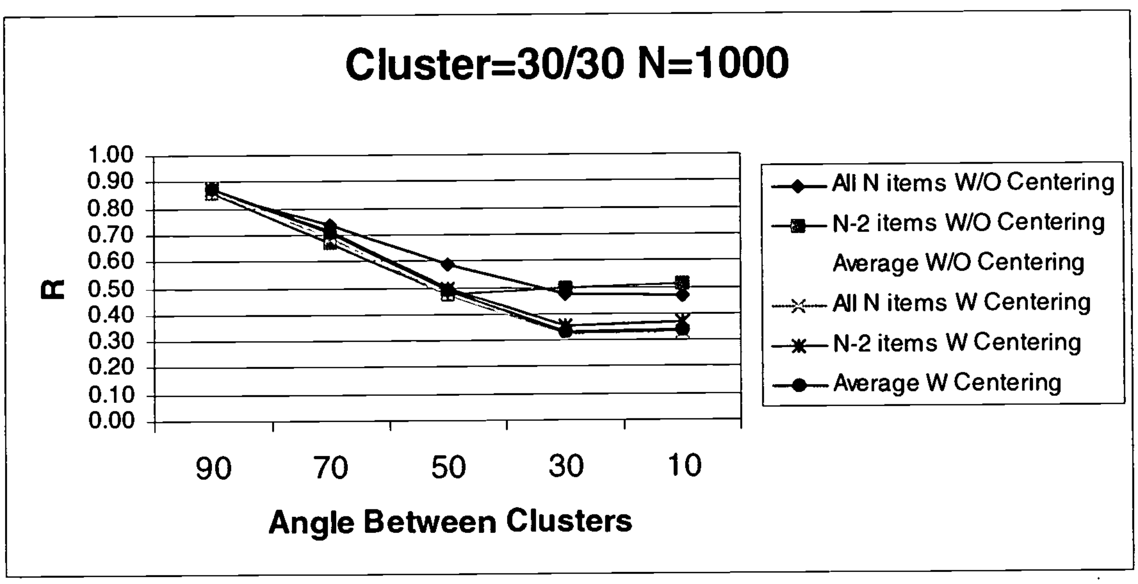


Figure 8d:



48

Figure 9: Average percentage of correct classification

Figure 9a:

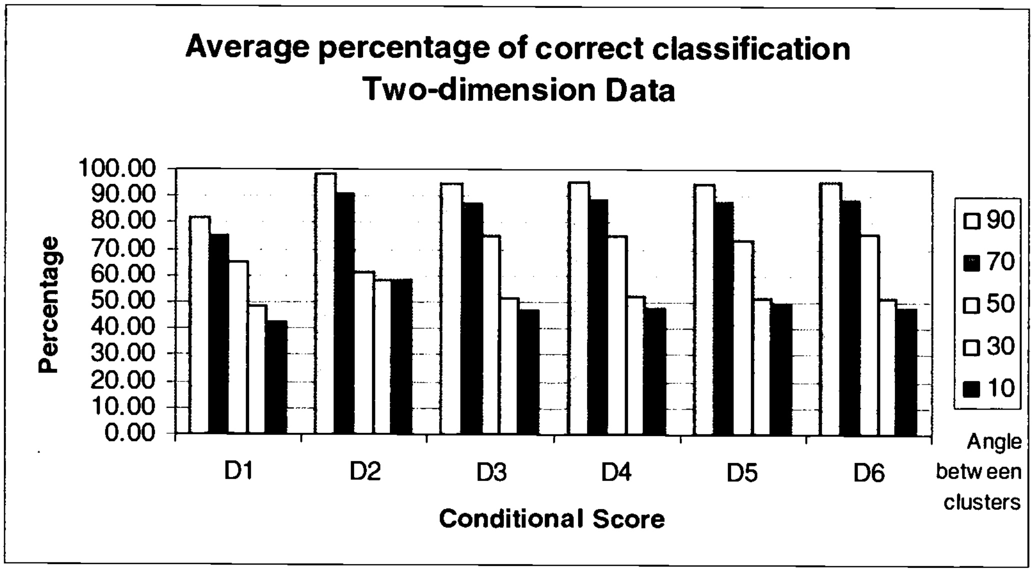
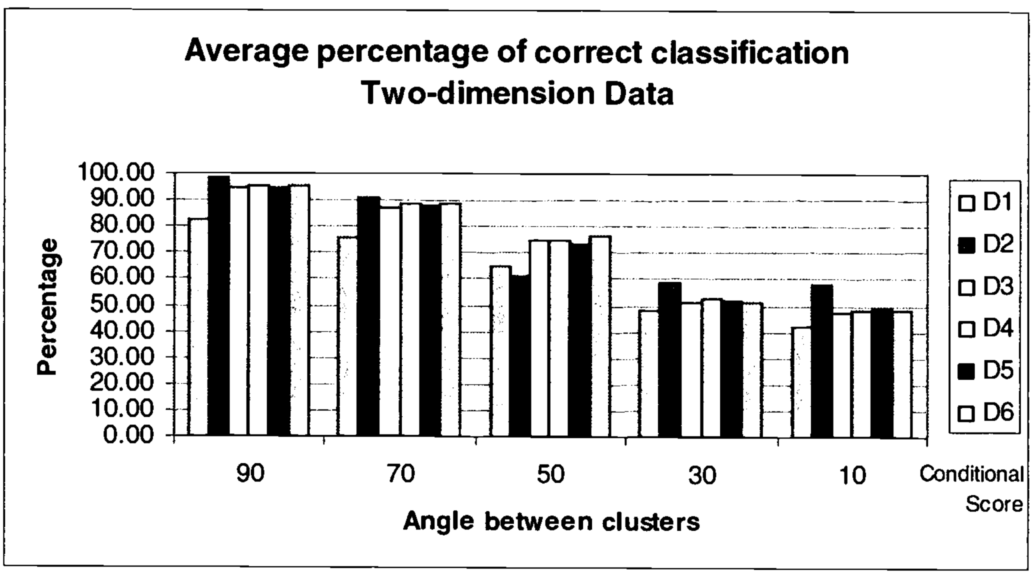


Figure 9b:



49

Figure 10: Average Dmax values

Figure 10a:

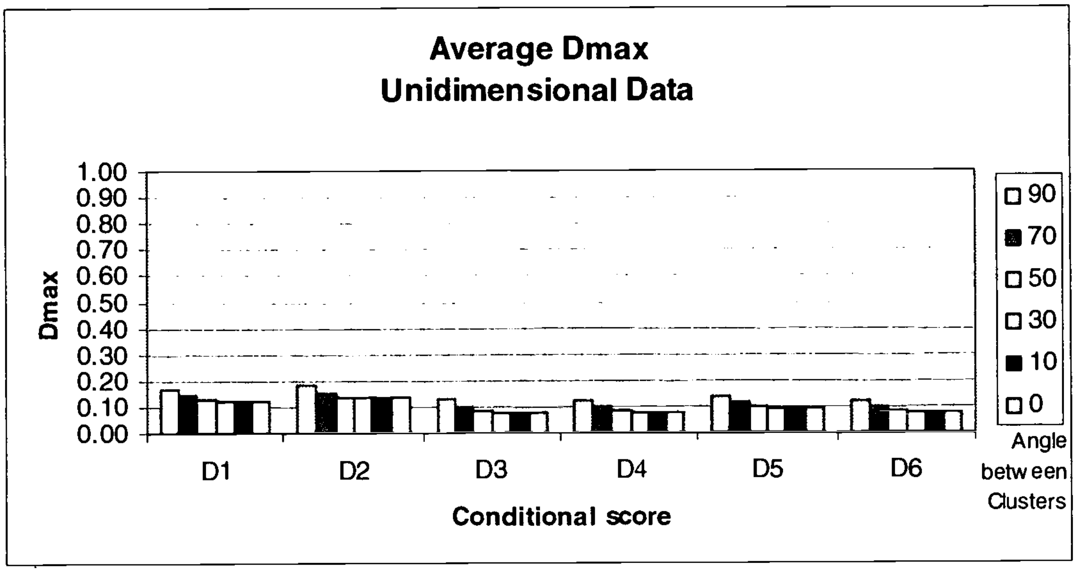
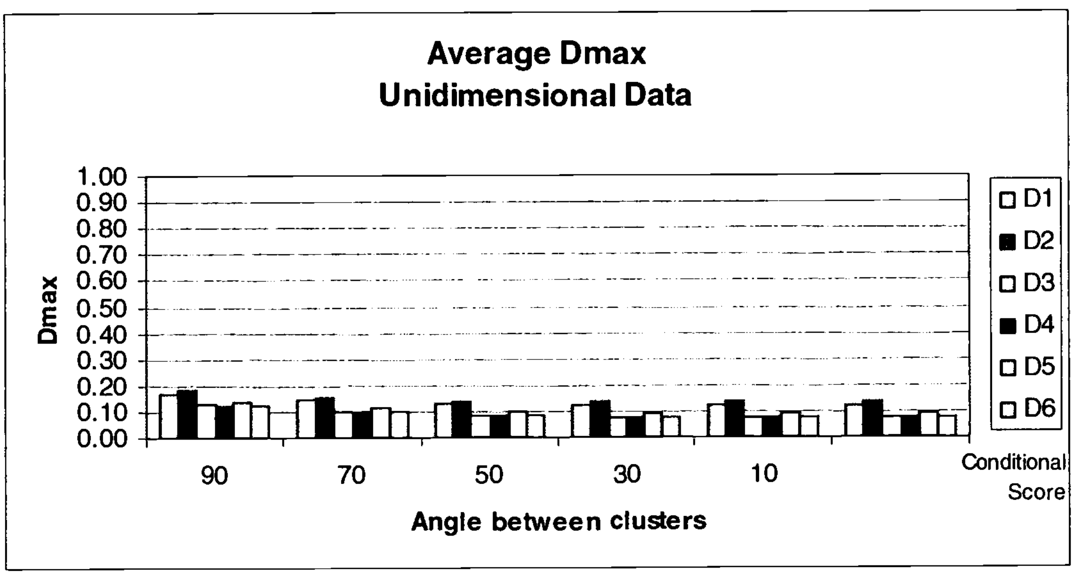


Figure 10b:



50

Figure 10c:

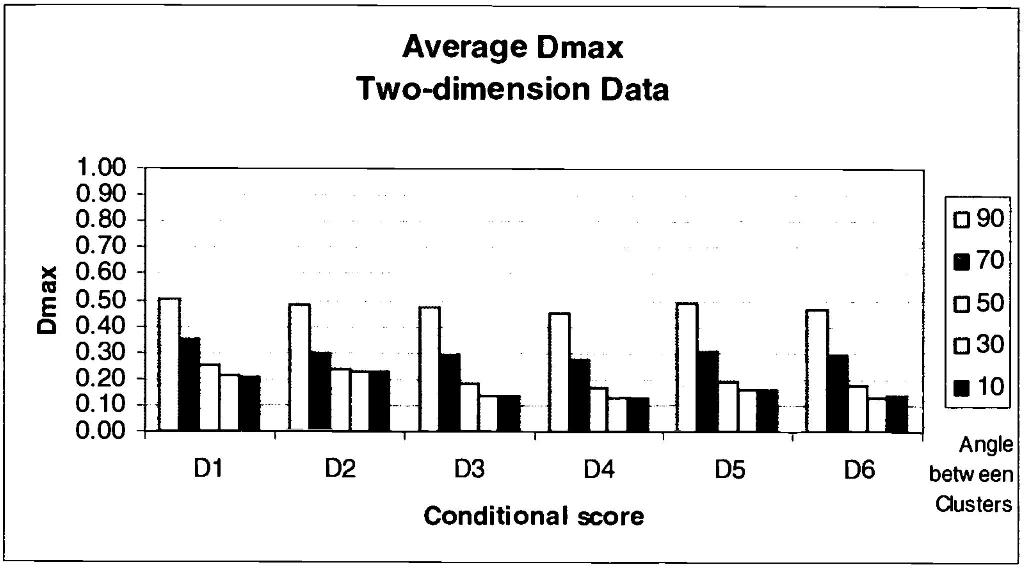
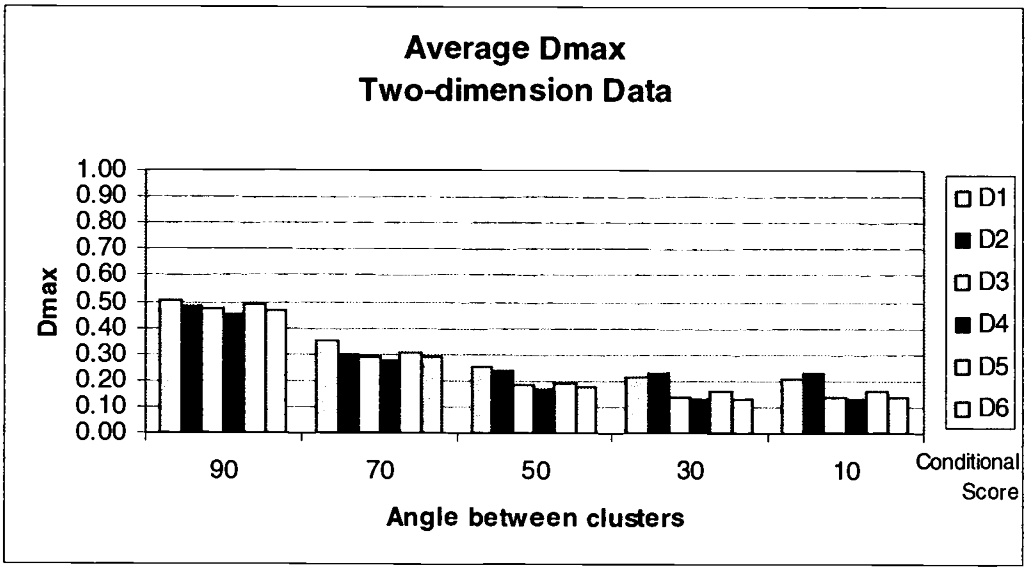


Figure 10d:



51

Figure 11: Average r ratio

Figure 11a:

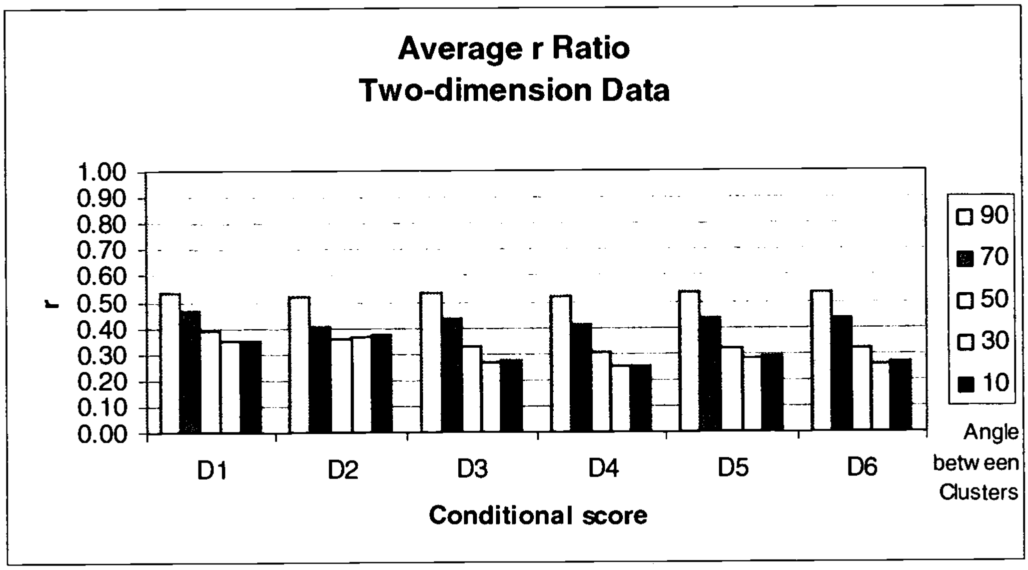
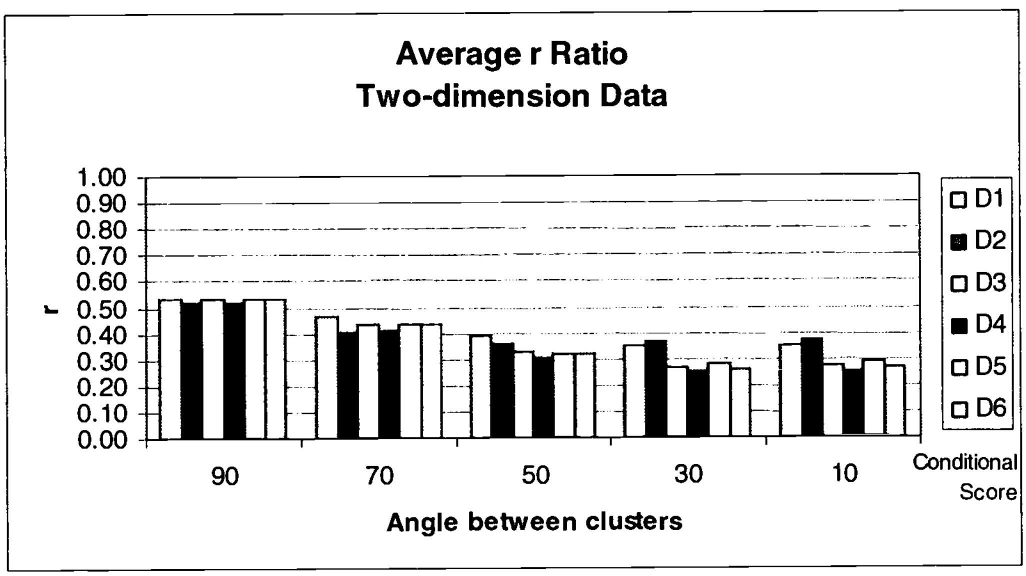


Figure 11b:



52

NCME 2003

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

ERIC
Educational Resources Information Center

REPRODUCTION RELEASE

(Specific Document)

TM035065

I. DOCUMENT IDENTIFICATION:

Title: The Impact of Conditional Scores on the Performance of DETECT

Author(s): Yanwei O. Zhang, Feng Yu, Ratna Nandakumar

Corporate Source: AICPA, ETS, University of Delaware

Publication Date: 2003

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic media, and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following three options and sign at the bottom of the page.

The sample sticker shown below will be affixed to all Level 1 documents	The sample sticker shown below will be affixed to all Level 2A documents	The sample sticker shown below will be affixed to all Level 2B documents
PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 1	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE, AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 2A	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 2B
Level 1	Level 2A	Level 2B
☑	☐	☐
Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g., electronic) *and* paper copy.	Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only	Check here for Level 2B release, permitting reproduction and dissemination in microfiche only

Documents will be processed as indicated provided reproduction quality permits.
If permission to reproduce is granted, but no box is checked, documents will be processed at Level 1.

I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.

Sign here, → please

Signature: Yanwei Zhang		Printed Name/Position/Title: YANWEI ZHANG	
Organization/Address: AICPA / 507 Bells Ct, Apt C2, Andalusia, PA 19020		Telephone: 215-633-7279	FAX:
		E-Mail Address: ozhang@aicpa.org	Date: 06/12/2003

(Over)

III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, *or*, if you wish ERIC to cite the availability of the document from another source, please provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents that cannot be made available through EDRS.)

Publisher/Distributor:
Address:
Price:

IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant this reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

Name:
Address:

V. WHERE TO SEND THIS FORM:

Send this form to the following ERIC Clearinghouse:

ERIC CLEARINGHOUSE ON ASSESSMENT AND EVALUATION
UNIVERSITY OF MARYLAND
1129 SHRIVER LAB
COLLEGE PARK, MD 20742-5701
ATTN: ACQUISITIONS

However, if solicited by the ERIC Facility, or if making an unsolicited contribution to ERIC, return this form (and the document being contributed) to:

ERIC Processing and Reference Facility
4483-A Forbes Boulevard
Lanham, Maryland 20706

Telephone: 301-552-4200
Toll Free: 800-799-3742
FAX: 301-552-4700
e-mail: ericfac@inet.ed.gov
WWW: http://ericfacility.org